

## PATENT COOPERATION TREATY

PCT

## NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents  
United States Patent and Trademark  
Office  
Box PCT  
Washington, D.C.20231  
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 25 August 2000 (25.08.00)	
International application No. PCT/US99/19068	Applicant's or agent's file reference 6523-020-228
International filing date (day/month/year) 18 August 1999 (18.08.99)	Priority date (day/month/year) 18 August 1998 (18.08.98)
Applicant XU, Tian et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:

16 March 2000 (16.03.00)

☐ in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was  
☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer Alejandro HENNING
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

# PATENT COOPERATION TREATY

From the  
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To: ADRIANE M. ANTLE  
PENNIE & EDMONDS LLP  
1155 AVENUE OF THE AMERICAS  
NEW YORK, NY 10036  
USA

## PCT

### NOTIFICATION OF TRANSMITTAL OF INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Rule 71.1)

Date of Mailing  
(day/month/year)

31 JAN 2001

Applicant's or agent's file reference

6523-020-228

#### IMPORTANT NOTIFICATION

International application No.

PCT/US99/19068

International filing date (day/month/year)

18 AUGUST 1999

Priority Date (day/month/year)

18 AUGUST 1998

Applicant

YALE UNIVERSITY

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.
2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.
3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.
4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices)(Article 39(1))(see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

SHIN-LIN CHEN

Telephone No. (703) 308-0196

## PATENT COOPERATION TREATY

## PCT

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

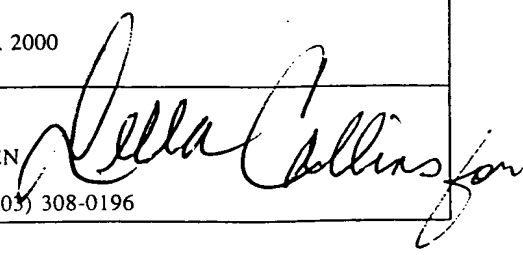
(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 6523-020-228	<b>FOR FURTHER ACTION</b> See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/US99/19068	International filing date (day/month/year) 18 AUGUST 1999	Priority date (day/month/year) 18 AUGUST 1998
International Patent Classification (IPC) or national classification and IPC Please See Supplemental Sheet.		
Applicant YALE UNIVERSITY		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 8 sheets.
- ☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).
- These annexes consist of a total of 0 sheets.

## 3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☒ Non-establishment of report with regard to novelty, inventive step or industrial applicability
- IV ☒ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☒ Certain observations on the international application

Date of submission of the demand 16 MARCH 2000	Date of completion of this report 28 DECEMBER 2000
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer SHIN-LIN CHEN 
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US99/19068

**I. Basis of the report**1. With regard to the **elements** of the international application:\*

- ☒ the international application as originally filed
- ☒ the description:  
pages 1-54 , as originally filed  
pages NONE , filed with the demand  
pages NONE , filed with the letter of \_\_\_\_\_
- ☒ the claims:  
pages 55-67 , as originally filed  
pages NONE , as amended (together with any statement) under Article 19  
pages NONE , filed with the demand  
pages NONE , filed with the letter of \_\_\_\_\_
- ☒ the drawings:  
pages 1-51 , as originally filed  
pages NONE , filed with the demand  
pages NONE , filed with the letter of \_\_\_\_\_
- ☒ the sequence listing part of the description:  
pages 1-37 , as originally filed  
pages NONE , filed with the demand  
pages NONE , filed with the letter of \_\_\_\_\_

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.  
These elements were available or furnished to this Authority in the following language \_\_\_\_\_ which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☒ contained in the international application in printed form.
- ☒ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. ☒ The amendments have resulted in the cancellation of:

- ☒ the description, pages NONE
- ☒ the claims, Nos. NONE
- ☒ the drawings, sheets/fig NONE

5. ☐ This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).\*\*

\* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

\*\*Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.



INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.  
PCT/US99/19068

**III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability**

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been and will not be examined in respect of:

☐ the entire international application.

☒ claims Nos. 24-67, 82, 83, 93, 99-113

because:

☐ the said international application, or the said claim Nos. \_ relate to the following subject matter which does not require international preliminary examination (*specify*).

☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. \_ are so unclear that no meaningful opinion could be formed (*specify*).

☐ the claims, or said claims Nos. \_ are so inadequately supported by the description that no meaningful opinion could be formed.

☒ no international search report has been established for said claims Nos. (See Attached).

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

☐ the written form has not been furnished or does not comply with the standard.

☐ the computer readable form has not been furnished or does not comply with the standard.

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US99/19068

**IV. Lack of unity of invention**

1. In response to the invitation to restrict or pay additional fees the applicant has:

- ☐ restricted the claims.
- ☒ paid additional fees.
- ☐ paid additional fees under protest.
- ☐ neither restricted nor paid additional fees.

2. ☐ This Authority found that the requirement of unity of invention is not complied with and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

- ☐ complied with.
- ☒ not complied with for the following reasons:

Please See Supplemental Sheet.

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

- ☐ all parts.
- ☐ the parts relating to claims Nos. . . .

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement****1. statement**

Novelty (N)	Claims	<u>(Please See supplemental sheet)</u>	YES
	Claims	<u>(Please See supplemental sheet)</u>	NO
Inventive Step (IS)	Claims	<u>(Please See supplemental sheet)</u>	YES
	Claims	<u>(Please See supplemental sheet)</u>	NO
Industrial Applicability (IA)	Claims	<u>(Please See supplemental sheet)</u>	YES
	Claims	<u>(Please See supplemental sheet)</u>	NO

**2. citations and explanations (Rule 70.7)**

Claims 1, 3, 5, 6, 8, 10-16, 18, 20 and 22 lack an inventive step under PCT Article 33(3) as being obvious over Yale University, 1996 (A) in view of Ligand Pharmaceuticals, Inc. (B), 1995.

Reference A teaches that lats gene encodes a protein which acts as a tumor suppressor to inhibit cell proliferation. Reference A teaches generation of a non-human animal comprising an inactivated lats gene via homologous recombination of nucleic acid comprising a non-lats sequence flanked by lats genomic sequence (e.g. p. 162, 163), and use of said non-human knockout animal as an animal model for studying diseases and disorders involving cell overproliferation (e.g. malignancy) and screening for test molecule having the ability to inhibit overproliferation (e.g. tumor formation) and to treat or prevent such diseases or disorders (e.g. p. 78, 79). Reference A also teaches a method of treating or preventing a disease or disorder involving cell proliferation, such as breast cancer, colon cancer, bladder cancer sarcoma, melanoma etc., in a subject by administering a therapeutically effective amount of a molecule that promote lats function (e.g. p. 158).

Reference A does not specifically teaches comparing the size of a cancer in a recombinant non-human animal with or without the administration of the potential therapeutic compound.

Reference B teaches a method for screening for a therapeutic agent for treatment of a pathological condition affected by the level of a cytokine by measuring the binding of a protein to a promoter, wherein a reduction in the binding in the presence of said agent compared to the binding in the absence of said agent is indicative of said agent as potentially useful for treatment of said condition.

It would have been obvious for one of ordinary skill at the time of the invention to use the non-human knockout animal as an animal model for screening agent as taught by A and compare the size of the cancer or tumor in said knockout animal in the presence or absence of said agent in order to identify a therapeutic compound that could promote lats function such as to provide therapeutic effects for the treatment of diseases or disorders involving cell overproliferation. such as cancers.

(Continued on Supplemental Sheet.)

**VIII. Certain observations on the international application**

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

The description is objected to under PCT Rule 66.2(a)(v) as lacking clarity under PCT Article 5 because it fails to adequately enable practice of the claimed invention because: the description only discloses a knockout mouse lacking functional lats gene (lats<sup>-/-</sup>) and fails to provide adequate guidance and evidence for any other non-human knockout animal lacking functional lats gene. The art in producing transgenic animal was unpredictable at the time the invention was made. One skilled in the art would not be able to predict the resulting phenotype of a non-human knockout lats<sup>-/-</sup> animal, and would require undue experimentation to practice over the full scope of the invention claimed.

Further, the description fails to provide adequate guidance and evidence for the therapeutic effects of a complex or a chimeric protein of a lats protein and a cdc2 protein. The art for gene therapy or protein therapy was unpredictable at the time of the invention. It would have required one skilled in the art undue experimentation to practice the claimed invention.

Claims 1-7, 10-17, 20-23, 89-92 and 94-98 are objected to as lacking clarity under PCT Rule 66.2(a)(v) because practice of the claimed invention is not enabled as required under PCT Rule 5.1(a) for the reasons set forth in the immediately preceding paragraph.

**Supplemental Box**

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

Sheet 10

**CLASSIFICATION:**

The International Patent Classification (IPC) and/or the National classification are as listed below:

IPC(7): A61K 39/395; C07H 21/04; A01N 61/00, 37/18, 43/04; C12N 5/00, 15/00 and US Cl.: 424/130.1; 536/23.5; 514/1, 2, 44; 800/3, 13, 14, 21, 25

**III. NON-ESTABLISHMENT OF REPORT:**

No international search report has been established for claim numbers 24-67, 82, 83, 93, 99-113.

**IV. LACK OF UNITY OF INVENTION:**

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2, and 13.3 is not complied with for the following reasons:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-23, drawn to recombinant non-human animal with knockout lats gene function, method of using said animal for screening compound in treating cancer.

Group II, claim(s) 68-77 and 89-91, drawn to a purified complex of a lats protein and cdc2 protein, and a pharmaceutical composition containing said complex.

Group III, claim(s) 78-81, 86, 88, 92, 95 and 97, drawn to a chimeric protein comprising fragment of lats protein and cdc2 protein, and a pharmaceutical composition containing said chimeric protein.

Group IV, claim(s) 84, 85, 87, 94, 96 and 98, drawn to a combination of nucleic acids comprising a nucleic acid encoding a lats protein and a nucleic acid encoding a cdc2 protein, and a pharmaceutical composition containing said nucleic acid.

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Transgenic animals, nucleic acids and proteins are different products which differ in physical properties, chemical structure and utilities, and do not share any common special technical feature. Groups I, II-III and IV do not share special technical features.

A chimeric protein comprising fragment of a lats protein and a cdc2 protein differs from a complex of protein comprising a lats protein and a cdc2 protein structurally and chemically. They are different products with different functions and different usages. Groups II and III do not share special technical features. Thus, Groups I-IV do not relate to a single inventive concept.

**V. 1. REASONED STATEMENTS:**

The report as to Novelty was positive (YES) with respect to claims 1-23, 68-81, 84-92, 94-98.

The report as to Novelty was negative (NO) with respect to claims NONE.

The report as to Inventive Step was positive (YES) with respect to claims 2, 4, 7, 9, 17, 19, 21, 23, 68-81, 84-92, 94-98.

The report as to Inventive Step was negative (NO) with respect to claims 1, 3, 5, 6, 8, 10-16, 18, 20, 22.

The report as to Industrial Applicability was positive (YES) with respect to claims 1-23, 68-81, 84-92, 94-98.

The report as to Industrial Applicability was negative (NO) with respect to claims NONE.

**V. 2. REASONED STATEMENTS - CITATIONS AND EXPLANATIONS (Continued):**

Claims 1-23, 68-81, 84-92 and 94-98 meet the criteria set out in PCT Article 33(2) and (4), because the prior art does not

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US99/19068

**Supplemental Box**

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

Sheet 11

teach or fairly suggest the combination of lats protein or nucleic acid with cdc2 protein or nucleic acid and the combination of lats and cdc2.

----- NEW CITATIONS -----

NONE

# PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To: ADRIANE M. ANTLER  
PENNIE & EDMONDS LLP  
1155 AVENUE OF THE AMERICAS  
NEW YORK, NY 10036  
USA

REFERRED TO  
REC'D

DEC 20 1999

Pennie & Edmonds  
O.K. for filing

## PCT

### NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT OR THE DECLARATION

(PCT Rule 44.1)

Date of Mailing  
(day/month/year) **16 DEC 1999**

Applicant's or agent's file reference  
6523-020-228

FOR FURTHER ACTION See paragraphs 1 and 4 below

International application No.  
PCT/US99/19068

International filing date  
(day/month/year)  
18 AUGUST 1999

Applicant  
YALE UNIVERSITY

1. ☒ The applicant is hereby notified that the international search report has been established and is transmitted herewith.

**Filing of amendments and statement under Article 19:**

The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

**When?** The time limit for filing such amendments is normally 2 months from the date of transmittal of the international search report; however, for more details, see the notes on the accompanying sheet.

**Where?** Directly to the International Bureau of WIPO  
34, chemin des Colombettes  
1211 Geneva 20, Switzerland  
Facsimile No.: (41-22) 740.14.35

For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. ☐ With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

- ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.  
☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Further action(s):** The applicant is reminded of the following:

Shortly after 18 months from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in rules 90 bis 1 and 90 bis 3, respectively, before the completion of the technical preparations for international publication.

Within 19 months from the priority date, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later).

Within 20 months from the priority date, the applicant must perform the prescribed acts for entry into the national phase before all designated Offices which have not been elected in the demand or in a later election within 19 months from the priority date or could not be elected because they are not bound by Chapter II.

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231  
Facsimile No. (703) 305-3230

Authorized officer

SHIN-LIN CHEN

Telephone No. (703) 308-0196

## PATENT COOPERATION TREATY

## PCT

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 6523-020-228	FOR FURTHER ACTION	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
International application No. PCT/US99/19068	International filing date (day/month/year) 18 AUGUST 1999	(Earliest) Priority Date (day/month/year) 18 AUGUST 1998
Applicant YALE UNIVERSITY		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 5 sheets.

☒ It is also accompanied by a copy of each prior art document cited in this report.

1. ☐ Certain claims were found unsearchable (See Box I).
2. ☒ Unity of invention is lacking (See Box II).
3. ☒ The international application contains disclosure of a nucleotide and/or amino acid sequence listing and the international search was carried out on the basis of the sequence listing

- ☒ filed with the international application.
- ☐ furnished by the applicant separately from the international application,  
☐ but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.
- ☐ transcribed by this Authority.

4. With regard to the title, ☐ the text is approved as submitted by the applicant.  
☒ the text has been established by this Authority to read as follows:

LATS KNOCK-OUT ANIMAL MODELS AND THEIR USES

5. With regard to the abstract,

- ☒ the text is approved as submitted by the applicant.
- ☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the drawings to be published with the abstract is:

Figure No. \_\_\_\_\_

- ☐ as suggested by the applicant.
- ☐ because the applicant failed to suggest a figure.
- ☐ because this figure better characterizes the invention.
- ☒ None of the figures.



# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/19068

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/19068

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : 424/130.1; 536/23.5; 514/1, 2, 44; 800/3, 13, 14, 21, 25

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/130.1; 536/23.5; 514/1, 2, 44; 800/3, 13, 14, 21, 25

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
WEST, STN, MEDLINE, CAPLUS, BIOSIS, SCISEARCH

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96/30402 A1 (YALE UNIVERSITY) 03 October 1996, entire document, especially pages 158-163.	1-5
Y		6-113
Y	WO 95/31722 A1 (LIGAND PHARMACEUTICALS, INC.) 23 November 1995, entire document, especially page 50.	6-37
Y, P		68-113
	TAO et al. Human homologue of the Drosophila melanogaster lats tumor suppressor modulates CDC2 activity, Nature Genetics, February 1999, Vol 21, No. 2, page 177-181, entire document.	

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search  
27 OCTOBER 1999

Date of mailing of the international search report

16 DEC 1999

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Authorized officer

SHIN-LIN CHEN

Facsimile No. (703) 305-3230

Telephone No. (703) 308-0196

**INTERNATIONAL SEARCH REPORT****International application No.**  
**PCT/US99/19068****C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

<b>Category*</b>	<b>Citation of document, with indication, where appropriate, of the relevant passages</b>	<b>Relevant to claim No.</b>
<b>A, P</b>	<b>ST. JOHN et al. Mice Deficient of Lats1 Develop Soft-Tissue Sarcomas, Ovarian Tumors and Pituitary Dysfunction, Nature Genetics, February 1999, Vol 21, No. 2, page 182-186, entire document.</b>	<b>1-37</b>

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/19068

## A. CLASSIFICATION OF SUBJECT MATTER: IPC (6):

A61K 39/395; C07H 21/04; A01N 61/00, 37/18, 43/04; C12N 5/00, 15/00

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

mtr

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>A61K 39/395, C07H 21/04, A01N 61/00, 37/18, 43/04, C12N 5/00, 15/00</b>		<b>A1</b>	(11) International Publication Number: <b>WO 00/10602</b> (43) International Publication Date: <b>2 March 2000 (02.03.00)</b>
(21) International Application Number: <b>PCT/US99/19068</b> (22) International Filing Date: <b>18 August 1999 (18.08.99)</b> (30) Priority Data: 60/096,996 18 August 1998 (18.08.98) US 60/096,997 18 August 1998 (18.08.98) US (71) Applicant (for all designated States except US): <b>YALE UNIVERSITY [US/US]; 451 College Street, New Haven, CT 06520 (US).</b> (72) Inventors; and (75) Inventors/Applicants (for US only): <b>XU, Tian [US/US]; 149 Flag Marsh Road, Guilford, CT 06437 (US). TAO, Wu-fan [US/US]; 420 Delaware Street S.E., Box 806 Mayo, Minneapolis, MN 55455 (US). ST. JOHN, Maie, A., R. [US/US]; 2351 Camden Avenue, Los Angeles, CA 90064 (US). FEI, Xiaolan [US/US]; Apartment 21, 393 Mansfield Street, New Haven, CT 06511 (US). FUKUMOTO, Royd, K. [US/US]; P.O. Box 1266, Alea, HI 96701 (US). ZHANG, Sheng [US/US]; 146 Cottage Street S.E., New Haven, CT 06511 (US). TURENCHALK, Gregory, S. [US/US]; 820 Orange Street, New Haven, CT 06511 (US). STEWART, Rodney, A. [US/US]; Apartment B, 512 Whitney Avenue, New Haven, CT 06511 (US).</b>			(74) Agents: <b>ANTLER, Adriane, M. et al.; Pennie &amp; Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).</b>  (81) Designated States: <b>AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</b>

**Published***With international search report.***(54) Title: LATS KNOCK-OUT ANIMAL MODELS AND THEIR USES****(57) Abstract**

A recombinant non-human animal having an inactivated *lats* gene is described. A *lats* knock-out mouse is exemplified. Because mice disrupted for the *lats* gene develop a variety of tumors, are susceptible to induction of skin tumors by exposure to carcinogens, and exhibit pituitary dysfunction, they have utility in screening for compounds effective to treat or prevent cancer or pituitary disorders. Compounds can be screened for activity in treating or preventing skin cancer in recombinant non-human animals which have an inactivated and in which skin tumors have been induced by exposure to carcinogens. Methods for treatment of cancers refractory to treatment with chemotherapy or radiation therapy by using a therapeutic that promotes *lats* function are also described. Additional methods are described for the treatment or prevention of diseases and disorders associated with aberrant levels of *cdc2* activity with a therapeutic that either promotes, inhibits or antagonizes *lats* function.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CN	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## LATS KNOCK-OUT ANIMAL MODELS AND THEIR USES

5           This invention was made with Government support under Grant number NIH-NCI 1  
R01 CA 69408 awarded by the National Institutes of Health. The Government has certain  
rights in the invention.

RELATED APPLICATIONS

10           Priority is claimed to United States provisional application Serial Nos. 60/096,997  
and 60/096,996, both filed on August 18, 1998, both of which are incorporated by reference  
herein in their entireties.

FIELD OF THE INVENTION

15           The present invention relates to the use of *lats* proteins, derivatives and fragments  
for the treatment of cancer, particularly for the treatment of cancer that is refractory to  
treatment by standard chemotherapy and radiation therapy protocols. The present invention  
also relates to the use of *lats* proteins, derivatives and fragments for the treatment of  
diseases and disorders associated with an aberrantly high or aberrantly low level of *cdc2*  
20 activity. The present invention further provides complexes of *lats* and *cdc2*, and their  
production and uses. The present invention also provides an animal model for cancer,  
particularly for skin cancer, soft tissue sarcomas, and ovarian tumors, and for pituitary  
disorders. The animal model is preferably a mouse, in which a *lats* gene has been disrupted  
by homologous recombination, *e.g.*, a *lats* knock-out mouse. The present invention also  
25 provides methods of screening potential therapeutics for efficacy in the treatment and  
prevention of cancer and pituitary disorders using *lats* knock-out animals.

BACKGROUND OF THE INVENTION**Cancer**

30           A neoplasm, or tumor, is a neoplastic mass resulting from abnormal uncontrolled  
cell growth, which may cause swelling on the body surface, and which can be benign or  
malignant. Benign tumors generally remain localized. Malignant tumors are collectively  
termed cancers. The term "malignant" generally means that the tumor can invade and  
destroy neighboring body structures and spread to distant sites to cause death (for review,  
35 see Robbins and Angell, 1976, *Basic Pathology*, 2d Ed., W.B. Saunders Co., Philadelphia,  
pp. 68-122). Treatment options, such as surgery, chemotherapy and radiation treatment, are

either ineffective or present serious side effects. Thus, there is a need for development of new drugs for the treatment of cancer.

### The Cell Cycle and Tumor Suppressors

5 Many cancers have been linked to perturbations in the regulation of the cell cycle, resulting in deregulation of cell growth. Briefly, the cell cycle occurs in four stages: G1 (for Gap1), the resting stage prior to DNA synthesis; S (for synthesis) phase, in which DNA synthesis occurs; G2 (for Gap2), the resting stage after DNA synthesis and prior to mitosis; and M phase, mitosis, in which cell division occurs. Progression of the cell cycle is driven  
10 by a group of cyclin-dependent kinases (CDKs) (Elledge, 1996, Science 274:1664-1672; Nasmyth, 1996, Science 274:1643-1645). The kinase activities of CDKs require their positive subunits, the cyclins, and the activities of specific CDK/cyclin complexes are in turn positively and negatively regulated by phosphorylation events and CDK inhibitors (CKIs) (Hunter and Pines, 1995, Cell 80:225-236; Morgan, 1995, Nature 374:131-134).  
15 While the specific CDKs, CDK2, CDK4 and CDK6, along with Cyclins D and E, regulate the progression from G1 into S phase, cdc2, along with Cyclins A and B, regulate the cell cycle progression from G1 into mitosis (Hunter and Pines, 1995, Cell 80:225-236).

Human tumor suppressors often act as negative regulators of the cell cycle, and several tumor suppressors are known to affect the activities of the CDK/cyclin complexes.  
20 For example, p53 activates the transcription of the p21 (p21<sup>WAF1/CIP1</sup>) CDK inhibitor in response to DNA damage signals, and p21 in turn binds and inactivates the CDK4 and CDK6 cyclin D complexes (Gartel et al., 1996, Proc. Soc. Exp. Biol. Med. 213:138-149). Another CDK inhibitor, p16, is itself a potent tumor suppressor (Biggs and Kraft, 1995, J. Mol. Med. 73:509-514). Although multiple members of the p16 and p21 inhibitor families have been identified for other major CDKs, corresponding inhibitors that regulate the  
25 mitotic CDK, cdc2, have not previously been identified (Morgan, 1995, Nature 374:131-134).

### Cancer Therapy

Currently, cancer therapy may involve surgery, chemotherapy and/or radiation  
30 treatment to eradicate neoplastic cells in a patient (see, for example, Stockdale, 1998, "Principles of Cancer Patient Management", in Scientific American: Medicine, vol. 3, Rubenstein and Federman, eds., Chapter 12, Section IV). All of these approaches pose significant drawbacks for the patient. Surgery, for example, may be contraindicated due to the health of the patient or may be unacceptable to the patient. Additionally, surgery may  
35 not completely remove the neoplastic tissue. Radiation therapy is only effective when the neoplastic tissue exhibits a higher sensitivity to radiation than normal tissue, and radiation therapy can also often elicit serious side effects.



With respect to chemotherapy, there are a variety of chemotherapeutic agents available for treatment of neoplastic disease. A significant majority of cancer chemotherapeutics act by inhibiting DNA synthesis, either directly, or indirectly by inhibiting the biosynthesis of the deoxyribonucleotide triphosphate precursors, to prevent DNA replication and concomitant cell division (see, for example, Gilman et al., Goodman and Gilman's: The Pharmacological Basis of Therapeutics, Eighth Ed. (Pergamon Press, New York, 1990)). These agents, which include alkylating agents, such as nitrosourea, anti-metabolites, such as methotrexate and hydroxyurea, and other agents, such as etoposides, campathecins, bleomycin, doxorubicin, daunorubicin, etc., although not necessarily cell cycle specific, kill cells during S phase because of their effect on DNA replication. Other agents, specifically colchicine and the vinca alkaloids, such as vinblastine and vincristine, interfere with microtubule assembly resulting in mitotic arrest. Chemotherapy protocols generally involve administration of a combination of chemotherapeutic agents to increase the efficacy of treatment.

Despite the availability of a variety of chemotherapeutic agents, chemotherapy has many drawbacks (see, for example, Stockdale, 1998, "Principles Of Cancer Patient Management" in Scientific American Medicine, vol. 3, Rubenstein and Federman, eds., ch. 12, sect. 10). Almost all chemotherapeutic agents are toxic, and chemotherapy causes significant, and often dangerous, side effects, including severe nausea, bone marrow depression, immunosuppression, etc. Additionally, even with administration of combinations of chemotherapeutic agents, many tumor cells are resistant or develop resistance to the chemotherapeutic agents. In fact, those cells resistant to the particular chemotherapeutic agents used in the treatment protocol often prove to be resistant to other drugs, even those agents that act by mechanisms different from the mechanisms of action of the drugs used in the specific treatment; this phenomenon is termed pleiotropic drug or multidrug resistance. Thus, because of drug resistance, many cancers prove refractory to standard chemotherapeutic treatment protocols. There is a significant need for alternative cancer treatments, particularly for treatment of cancer that has proved refractory to standard cancer treatments, such as surgery, radiation therapy, and chemotherapy.

### **Pituitary Disorders**

The pituitary regulates numerous biological functions through its secretion of different hormones. (For review see Frohman, "The Anterior Pituitary" in Cecil Textbook of Medicine, 18th Ed., Wyngaarden and Smith, eds. (W.B. Saunders Company, Philadelphia, 1988) pp. 1290-1305). In particular, the pituitary releases glycoprotein hormones, which include luteinizing hormone (LH) and follicle stimulating hormone (FSH); LH and FSH regulate ovarian and testicular development as well as reproductive functions such as ovulation and spermatogenesis. Disruption of LH or FSH secretion has dramatic consequences for reproductive function, particularly for ovulation in the female. The pituitary also releases somatomammotropic hormones, including growth hormone and

prolactin. Growth hormone promotes linear growth and is involved in the regulation of certain metabolic functions such as sugar and amino acid uptake and use of fat stores. Prolactin stimulates and maintains lactation in post-parturition females. Although an increase or decrease in prolactin levels does not appear to have significant biological consequences beyond an effect on lactation, disruption of growth hormone secretion stunts growth and has other metabolic effects. Other pituitary hormones include corticotropin (ACTH), thyroid stimulating hormone (TSH), and endorphins and related peptides. Although in some situations, hormone replacement therapy is available, there is a need for additional therapeutics to treat or prevent pituitary dysfunctions.

## 10 LATS

The *large tumor suppressor* or *lats* gene (also known as *warts*), a tumor suppressor gene, was previously isolated from *Drosophila* using a mosaic screen. Inactivation of *lats* in somatic cells causes dramatic overproliferation phenotypes (Xu et al., 1995, Development 121:1053-1063; Justice et al., 1995, Genes & Devel. 9:534-546). Somatic cells that are mutant for *lats* undergo extensive proliferation and form large tumors in many tissues of mosaic flies (Xu et al., 1995, Development 121:1053-1063). Tumors that result from inactivation of *lats* display many features of human neoplasms. *Lats* mutant cells grow aggressively, and a single mutant cell can develop into a tumor that is 1/5 the size of the animal, and these fly tumors are highly irregular in shape and size and are often poorly differentiated (St. John and Xu, 1997, Am. J. Hum. Genet 61:1006-1010). *Drosophila* that are homozygous for the various *lats* alleles display a wide range of developmental defects including embryonic lethality, overproliferation of both neural and epidermal tissues, rough eyes, and sterility. Molecular characterization of *lats* indicates that it contains a putative kinase domain (Xu et al., 1995, Development 121:1053-1063; Justice et al., 1995, Genes & Devel. 9:534-546).

Mouse and human homologs of the *Drosophila lats* have also been identified, and human *lats* was found to be down-regulated in a large number of human tumor cell lines. The nucleotide and amino acid sequences of human *lats* (h-*lats*), mouse *lats* (m-*lats*), mouse *lats2* (m-*lats2*) and *Drosophila lats* are provided herein in Figures 12-15, respectively (SEQ ID NOS:1-8, respectively), and are described in PCT Publication WO 96/30402, published October 3, 1996, which is incorporated by reference herein in its entirety.

Citation of references hereinabove shall not be construed as an admission that such references are prior art to the present invention.

## SUMMARY OF THE INVENTION

The present invention relates to therapeutic and prophylactic methods and compositions for the treatment and prevention of cancers based on *lats* proteins, and therapeutically and/or prophylactically effective analogs and fragments of *lats* protein. This is due to the fact that although most tumor suppressor genes regulate the G1/S phase of the

cell cycle, the *lats* protein interacts with the cell cycle-dependent kinase *cdc2*, which is involved in the regulation of the G2 to M transition of the cell cycle, and thereby provides a means to regulate the G2 to M transition of the cell cycle and to treat cancers that have proven refractory to other cancer treatments, including chemotherapy and radiation therapy treatments. The invention provides for treatment and prevention of cancer by  
5 administration of a therapeutic compound of the invention. The therapeutic compounds of the invention useful for treatment of cancer refractory to a chemotherapy and/or radiation therapy protocol include: *lats* proteins, and therapeutically effective analogs and derivatives (including fragments) of *lats*, nucleic acids encoding *lats* proteins and therapeutically effective analogs and derivatives of *lats*, and *lats* agonists.

10 The invention further provides assays, both *in vivo* and *in vitro*, for testing the efficacy of the therapeutics of the invention for treatment of cancer, particularly cancer that has been shown to be refractory to chemotherapy and radiation therapy treatments.

In another aspect, the invention provides compositions and methods of production of complexes of *lats* and *cdc2* proteins ("lats-*cdc2* complexes"), including complexes of *lats*  
15 analogs or derivatives and *cdc2* analogs and derivatives (including complexes of *lats* proteins with *cdc2* analogs and derivatives and vice versa), where the analogs and derivatives have the ability to interact with the other member of the complex.

The phosphorylated form of *lats* complexes with *cdc2*. Accordingly, in a preferred embodiment, the *lats*-*cdc2* complexes contain phosphorylated *lats* protein, specifically *lats* protein phosphorylated on a serine or threonine residue within 20 residues upstream of an  
20 Ala-Pro-Glu consensus in subdomain eight of a *lats* kinase domain, *e.g.*, corresponding to serine 909 of human *lats*, as depicted in Figure 12 (SEQ ID NO:2). Alternatively, the *lats* protein in the *lats*-*cdc2* complex has a glutamate or aspartate residue substituted for a serine or threonine residue within 20 residues upstream of an Ala-Pro-Glu consensus in subdomain eight of a *lats* kinase domain, *e.g.*, corresponding to serine 909 of human *lats*, as depicted in Figure 12 (SEQ ID NO:2).

25 In another embodiment, the *lats*-*cdc2* complex contains a portion of *lats* protein corresponding to amino acids 15-585 of human *lats*, as depicted in Figure 12 (SEQ ID NO:2).

The invention further provides methods of modulating the activity of *cdc2* using *lats* proteins, as well as *lats* derivatives and fragments able to interact with *cdc2* protein, *lats*-  
30 *cdc2* complexes, and antibodies against *lats*-*cdc2* complexes. In particular, the invention provides methods for treating or preventing disorders involving an aberrant level of *cdc2* in a subject. Therapeutically effective amounts of compounds are administered to promote or inhibit LATS function, as required.

The invention provides recombinant non-human animals in which a *lats* gene has been inactivated, preferably recombinant mice in which a *lats* gene (preferably a gene  
35 having *lats* coding sequence of SEQ ID NO:3) has been inactivated, *i.e.*, a *lats* knock-out mouse.

In a preferred embodiment, the invention provides a *lats* knock-out mouse in which the inactivated *lats* gene had the coding sequence of SEQ ID NO:3, prior to disruption, and in a more preferred embodiment, the inactivated *lats* gene is deleted for the Lats C-terminal domain 1 (LCD1), the Lats C-terminal domain 2 (LCD2), the Lats C-terminal domain 3 (LCD3), and all or a portion of the kinase domain, and retains the Lats flanking domain (LFD), the Lats split domain 1 (LSD1), the Lats split domain 2 (LSD2), and the putative SH3-binding domain, in a most preferred embodiment the *lats* gene is disrupted by replacement of a non-*lats* sequence for the sequence encoding the amino acids corresponding to amino acids 756 to 1130 of human *lats*, as depicted in Figure 12 (SEQ ID NO:2). In other embodiments, the inactivated *lats* gene is deleted for all or a portion of the kinase domain (*e.g.*, so as to inactivate kinase activity). A *lats* "knock-out" animal is an animal in which at least one genomic copy of a *lats* gene has been inactivated by insertional mutagenesis, *e.g.*, by homologous recombination, for example, as described and exemplified herein.

The invention further provides methods for screening potential therapeutics for activity in the treatment or prevention of cancer, preferably soft tissue sarcomas and ovarian tumors, using the *lats* knock-out animals of the invention. The invention also provides methods for screening potential therapeutics for activity in the treatment or prevention of pituitary dysfunctions, using the *lats* knock-out animals of the invention. In a preferred embodiment, the invention provides methods for screening potential therapeutics for activity in the treatment or prevention of skin cancer using a non-human *lats* knock-out animal, preferably a *lats* knock-out mouse, in which skin tumors have been induced with carcinogens.

### Definitions

As used herein, underscoring or italicizing the name of a gene shall indicate the gene, in contrast to its encoded protein product which is indicated by the name of the gene in the absence of any underscoring or italicizing. For example, "*lats*" shall mean the *lats* gene, whereas "lats" shall indicate the protein product of the *lats* gene.

ARN = After Removal of Nocodazole

CDK = Cyclin Dependent Kinase

mlats = mouse *lats*

### BRIEF DESCRIPTION OF THE FIGURES

Figures 1A-H. Human *lats* can functionally replace the fly gene. (A) Adult *Drosophila* in which *lats* homozygous mutant cells have been induced in the imaginal tissues of the *lats* heterozygous larvae, exhibit *lats* mosaic phenotype, and have mutant cells which have undergone extensive proliferation and formed tumors in various body parts of the mosaic adults. (B) Adult *Drosophila* that express human *lats* (hs-h-*lats*) completely do not exhibit tumor development (compare to A). (C, D) High magnification views of the

flies in panel (B) show *yellow* bristles (white arrows), indicating *lats* cells (genetically marked by the *yellow* mutation) were produced and that they have developed into normal structures. (E, F) *lats<sup>e26-1</sup>* homozygous mutants display giant pupae and disc-overproliferation phenotypes (left pupae in panels). Without heat-shock induction, the leaky expression of the *hs-h-lats* transgene partially suppressed the *lats<sup>e26-1</sup>* phenotypes (middle pupae in panels). Daily induction of the *hs-h-lats* gene completely rescued the mutant phenotypes (right pupae in panels). (G) Scanning Election Micrograph view of a *lats* mosaic fly. (H) A *lats* tumor on the wing (indicated by an arrow in panel G) is enlarged, showing that cells in the overproliferated mutant clone have differentiated into wing cells with hair structures.

Figures 2A-E. Phosphorylation of *lats* oscillates with the cell cycle. (A) The phosphorylation of *lats* protein in HeLa cells after exposure of the cells to certain conditions was assayed by immunoprecipitation and blotted with an anti-h-*lats* monoclonal antibody. "CIP" indicates that the cells were incubated in calf intestinal phosphatase and "β-gp" indicates that the cells were incubated in β-glycerol phosphate. The "Time (min.)" indicates time in minutes of incubation. "⊖-h-Lats" and "h-Lats" indicate the phosphorylated and dephosphorylated forms of h-*lats*, respectively. *Lats* proteins from mitotic HeLa cell lysates (50 min. ARN (After Removal of Nocodazole)) display a slow-migrating form on SDS-PAGE (6%) (lane 1). The proteins are converted into a fast-migrating form when incubated with Calf Intestinal Phosphatase ("CIP") (lanes 2-4). When both CIP and a CIP inhibitor, β-glycerol phosphate ("β-gp"), are present, their mobility remains unchanged (lane 5). *Lats* proteins from 125 min ARN cells have both the slow-migrating and fast-migrating forms (lane 6) and CIP-treatment converts all *lats* proteins into the fast-migrating form (lanes 7-9). (B) Immunowestern blot shows that phosphorylation of the *lats* protein oscillates with the cell cycle. Cell cycle stages G0, G1, S, and G2, are indicated above each lane and cells in different mitotic stages (M) are indicated by min. (minutes ARN). The faint bands are degradation products of *lats*. The progression of the cell cycle was verified by DAPI staining. (C-E) These panels show fluorescent micrographs of DAPI staining for cells at three time points (50' (C), 75' (D), and 100' (E) ARN). Arrows indicate cells at metaphase (50' ARN (panel C)), anaphase (75' ARN (panel D)), or telophase (100' ARN (panel E)), respectively.

Figures 3A-E. *Lats* directly complexes with *cdc2* during mitosis and the *lats/cdc2* complex is inactive for H1 kinase activity. (A) *Cdc2* is co-immunoprecipitated with *lats* from mitotic CHO cell lysates (M) but not from quiescent CHO cell lysates (G0). Anti-h-*lats* polyclonal antibodies or anti-human cyclin B monoclonal antibodies were used for immunoprecipitation, and anti-human *cdc2* monoclonal antibodies were used to visualize *cdc2*. (B) *Cdc2* co-immunoprecipitated with human *lats* proteins at early mitosis. The stages of the cell cycle are indicated above each lane as "min. ARN" or "G0". *Lats* proteins were immunoprecipitated using anti-h-*lats* monoclonal antibodies and were separated on 8% SDS-PAGE. The western blot was sequentially probed with anti-human *cdc2*

monoclonal antibodies (lower panel, labeled "Cdc2"), anti-h-lats monoclonal antibodies (upper panel, labeled "h-lats"), anti-human cyclin B and cyclin A monoclonal antibodies (middle panels, labeled "Cyclin B" and "Cyclin A", respectively). (C) Co-immunoprecipitation of baculovirus-expressed human lats and cdc2 proteins. H-lats proteins were precipitated with anti-h-lats monoclonal antibodies and probed with anti-human cdc2 monoclonal antibodies (upper panel, labeled "Cdc 2") and cdc2 proteins were precipitated with anti-human cdc2 monoclonal antibodies and probed with anti-h-lats monoclonal antibodies (lower panel, labeled "h-lats"). (D) Lats-associated cdc2 is inactive for H1 kinase activity. Cdc2 proteins co-immunoprecipitated with either lats or cyclin B (indicated by "+" in the legend labeled "anti-Cyclin B" or "anti-h-lats") from 50 min. ARN HeLa cell lysates were divided for western quantification of cdc2 (upper panel, labeled "IP-Western") and for the histone H1 kinase assay (lower panel, labeled "H1 Kinase assay"), respectively. As a control, Protein G-agarose beads incubated with equal amounts of cell lysates only were also used for the H1 kinase assay (indicated by "-" in the legend). (E) Summarized are the results of yeast two-hybrid assays for interactions among full-length h-lats (h-lats), N-terminal region of h-lats (N-h-lats), C-terminal region of h-lats (C-h-lats), human cdc2, CDK2, and CDK4. +++, ++, +: indicate strong, intermediate, and weak interactions, respectively, while - and ND indicate no interaction and not determined, respectively.

Figures 4A-F. Genetic interaction between *lats*, *cdc2*, and *cyclin A* in *Drosophila*.

(A) *lats*<sup>P8</sup>/*lats*<sup>P8</sup> homozygotes die at the pupal stage. (B) Removal of one copy of the *cdc2* gene rescues *lats*<sup>P8</sup> lethality (*lats*<sup>P8</sup>/*lats*<sup>P8</sup>; +/*cdc2*<sup>B47</sup>). (C) A typical rough, overproliferated eye dissected from a *lats*<sup>P8</sup>/*lats*<sup>P8</sup> dead pupa. (D) An eye from the fly in panel (B), showing that the eye phenotype has been almost completely suppressed. (E) *lats*<sup>e26-1</sup>/*lats*<sup>e26-1</sup> pupae (the pupa labeled 1) are much larger than wild-type (the pupa labeled 4). The giant-pupa phenotype is partially reduced by a *cdc2* temperature sensitive mutant at room temperature (pupa labeled 2) (*lats*<sup>e26-1</sup>/*lats*<sup>e26-1</sup>; *cdc2*<sup>E1-24</sup>/*cdc2*<sup>E1-24</sup>) or by removal of one copy of the *cdc2* gene ([i]a labeled 3) (*lats*<sup>e26-1</sup>/*lats*<sup>e26-1</sup>; +/*cdc2*<sup>B47</sup>). (F) The third instar larval wing discs dissected from animals of the same genotypes in (E) and labeled with the same number. The *lats* disc overproliferation phenotype (larva 1) is dramatically suppressed by mutations in the *cdc2* gene (larvae 2 and 3). Removal of one copy of the *cyclin A* gene resulted in a phenotypic suppression of the *lats* mutants in a manner similar to removal of one *cdc2* gene as shown above.

Figures 5A-N. Effect of inactivation and overexpression of *lats* on the cell cycle in *Drosophila*. (A) *Drosophila* third instar eye imaginal disc contains a homozygous *lats*<sup>l</sup> clone (arrowhead indicates the lack of Myc staining) that crosses the morphogenetic furrow (MF) (arrow). (B) Cyclin A staining (indicated by arrowhead) in the clone exhibits expression that spans the MF (arrow). (C) Composite staining of the same disc shown in panels A and B showing Myc stains and propidium iodide staining which more clearly delineates the MF and the *lats* mutant clone (indicated by the arrow) that spans it. (D)

Composite staining magnification of the same *lats* mutant clone shown in panels A-C. Cyclin A staining clearly spans the MF region (arrow). (E) Composite staining of the same region shown in panel D viewed with Myc and propidium iodide. (F) Composite staining magnification of the same clone shown in panel E viewed with Cyclin A staining and propidium iodide. Cyclin A is degraded in *lats* mutant cells at late mitosis (as indicated by the arrowheads). (G) *Drosophila* third instar eye disc containing homozygous *lats<sup>xl</sup>* clones (lack of Myc staining is indicated by arrowheads) that span the MF region. (H) The third instar eye disc shown in panel G was also stained for Cyclin B. The MF (indicated by the arrow) is well defined and free of cyclin B staining. (I and J) Scanning electron micrograph of a wild-type *Drosophila* adult eye (I) in comparison to an eye (J) of a *GMR-d-lats* transgenic fly showing fewer and irregular ommatidia and missing bristles. (K) A section of a *GMR-d-lats* eye reveals that many pigment cells are missing, and that ommatidia occasionally lack a full complement of photoreceptor cells (arrow). (L) Propidium iodide staining of a *GMR-d-lats* third instar eye imaginal disc. A stripe of intensely stained nuclei which are tetraploid (indicated by the star) are seen in the region of the second mitotic wave, which is immediately followed posteriorly by apoptotic cells with fragmented nuclei (indicated by the small arrows). (M) BrdU labeling (green) of a *GMR-d-lats* third instar eye disc reveals that S phase occurs in the second mitotic wave (indicated by the star) posterior to the MF (indicated by the arrow) just as it would in wild type. (N) A *GMR-p21* eye disc is shown in which cell proliferation in the second mitotic wave is blocked before S phase and the stripe of BrdU labeling posterior to the MF (indicated by the arrow) is abolished.

Figures 6A-D. Targeting of the *lats* locus by homologous recombination. (A) Sequence alignment of human *lats* (h-*lats*) and mouse *lats* (m-*lats*, partial sequence). Arrow indicates the point at which the mouse *lats* gene was disrupted. (B) Targeting vector for positive-negative selection of homologous recombinants at the *lats* locus, with restriction map and the structure of the targeted *lats* locus. The vector is represented by the second line from the top, while the wild-type and mutant (i.e., disrupted) *lats* alleles are indicated by the top and bottom lines, respectively. The BamHI sites are indicated by "B", the EcoRI sites are indicated by "R", and the EcoRV sites are indicated by "RV". Exons are represented by filled rectangles. A BamHI/EcoRV double digest generates a 3.5 kb fragment from the wild-type allele and a 5.8 kb fragment from the disrupted allele, both of which are recognized by the probe shown, which is not contained in the targeting vector. In the vector and the mutant allele, the PGK-TK gene cassette and the PGK-neo fragment are denoted by open boxes labeled accordingly. (C) Southern blot of genomic DNA isolated from individual embryonic stem cell clones. The genotypes of the clones are indicated above the lanes with the "+/+" indicating wild-type clones, "+/-" indicating clones heterozygous for the mutant allele, and "-/-" indicating clones homozygous for the mutant allele. (D) Western blot using anti-h-*lats* polyclonal antibody on lysates from 13.5 dpc (days post coitus) mouse embryonic fibroblasts indicating the absence of *lats* protein in the knock-out mice. The genotype of the clones is indicated above the lanes as in panel C.

Figures 7A and B. Growth retardation of *lats*<sup>-/-</sup> mice. (A) Representative picture of a *lats*<sup>-/-</sup> mouse (agouti, the mouse on the right) with its wild-type littermate (black, the mouse on the left) (12-days-old). (B) Representative growth curve of *lats*<sup>+/+</sup>, *lats*<sup>+/-</sup>, and *lats*<sup>-/-</sup> mice. Mice were weighed (grams (g)) at intervals and plotted against age in days.

Figures 8A-D. Ovarian phenotypes of *lats*<sup>-/-</sup> mice. Histopathological sections of ovaries derived from *lats*<sup>+/+</sup> (panels A and C) and *lats*<sup>-/-</sup> (panels B and D) females. Overview images are in panels A and B while high magnification views are in panels C and D. The paraffin sections were stained with hematoxylin and eosin. An absence of corpora lutea (CL) is evident in the *lats*<sup>-/-</sup> ovary. Ovarian stromal cell tumors (SC) which obliterate the normal structure of the ovary, eliminating follicles (FC) progressively, are readily apparent.

Figures 9A-F. Absence of mammary gland development in *lats*<sup>-/-</sup> Mice. (A,B) *Lats*<sup>+/+</sup> female with normal mammary gland and nipple development. (D,E) *Lats*<sup>-/-</sup> female displaying absence of mammary gland and nipple formation. (C,F) Hematoxylin and eosin stained histopathological sections of mammary glands derived from *lats*<sup>+/+</sup> (C) and *lats*<sup>-/-</sup> (F) mice. The amount of breast epithelial tissue was markedly decreased in *lats*<sup>-/-</sup> females, resulting in mammary fat pads, devoid of an epithelial component.

Figures 10A-E. Pituitary hyperplasia and dysfunction in *lats*<sup>-/-</sup> mice. (A and B) Histopathological sections of pituitaries derived from *lats*<sup>+/+</sup> (A) and *lats*<sup>-/-</sup> (B) mice. The paraffin sections were stained with hematoxylin and eosin. The normal pituitary gland from a *lats*<sup>+/+</sup> mouse demonstrates the organized architecture of the gland. Hyperplastic changes are visible in *lats*<sup>-/-</sup> pituitaries. Multiple atypical cells showing irregularly shaped nuclei, and variability in shape and size are readily apparent. (C-E) The graphs indicate the amount of LH (C), PRL (D), and FSH (E) production in the mice. Hormone levels are plotted as ng/ml with the results from the *lats*<sup>+/+</sup> mice represented by the right bar of each bar graph and the results from the *lats*<sup>-/-</sup> mice represented by the left bar of each bar graph.

Figures 11A-C. Soft tissue sarcoma development in *lats*<sup>-/-</sup> mice. (A,B) Typical soft tissue sarcomas in *lats*<sup>-/-</sup> mice (A, 6.5 months old; B, 4.5 months old). (C) Histopathological section of the soft tissue sarcoma shown in panel B stained with hematoxylin and eosin revealing pleiomorphic, spindle-shaped cells characteristic of this tumor.

Figure 12. Nucleotide and amino acid sequences of human *lats* (h-*lats*) (SEQ ID NOS:1 and 2, respectively).

Figure 13. Nucleotide and amino acid sequences of mouse *lats* (m-*lats*) (SEQ ID NOS:3 and 4, respectively).

Figure 14. Nucleotide and amino acid sequences of mouse *lats2* (m-*lats2*) (SEQ ID NOS:5 and 6, respectively).

Figure 15. Nucleotide and amino acid sequences of *Drosophila lats* (SEQ ID NOS:7 and 8, respectively).



### DETAILED DESCRIPTION OF THE INVENTION

The present invention provides lats-cdc2 protein complexes, including complexes that contain lats analogs and fragments and/or cdc2 analogs and fragments, as well as methods of producing these complexes and nucleic acids encoding the two members of the complex. The invention also provides antibodies that bind immunospecifically to a lats-cdc2 complex, but do not bind the individual binding partners immunospecifically.

The invention also provides methods for the modulation of cdc2 activity using lats proteins and lats analogs and derivatives that are able to interact with cdc2. In particular, methods are provided for treating or preventing diseases and disorders associated with aberrant cdc2 activity by administration of a therapeutic compound of the invention.

The present invention further provides recombinant non-human animals, preferably mice, having at least one copy of (preferably both copies of, *i.e.*, is homozygous for) an inactivated *lats* gene, *i.e.*, *lats* knock-out animals. Preferably, these *lats* knock-out animals are generated by homologous recombination, *i.e.*, have a gene disrupted by insertional mutagenesis induced by homologous recombination with a nucleic acid containing non-lats sequences flanked by *lats* genomic sequences. The invention further provides methods of screening for compounds effective to treat or prevent cancer, preferably soft tissue sarcomas or ovarian tumors, more preferably skin cancer, using the recombinant non-human animals of the invention. The invention also provides methods of screening for compounds effective to treat or prevent pituitary dysfunction using the recombinant non-human animals of the invention.

Therapeutics of the invention that can be used to treat or prevent diseases and disorders associated with an aberrant level of cdc activity include those therapeutics that promote lats function (*e.g.*, lats proteins and lats derivatives and analogs that supply lats function, nucleic acids encoding lats, lats derivatives and analogs, and lats-cdc2 complexes), and those therapeutics that inhibit or antagonize lats function (*e.g.*, lats derivatives and analogs that inhibit or antagonize lats function) and nucleic acids encoding these lats derivatives and analogs, anti-lats antibodies and anti-lats-cdc2 complex antibodies, lats antisense nucleic acids, and lats inhibitors and antagonists.

The present invention also provides therapeutic methods and compositions for the treatment and prevention of cancer based on lats proteins and therapeutically or prophylactically effective analogs and fragments of lats proteins. The invention provides for treatment and prevention of cancer by administration of a therapeutic compound of the invention. The therapeutic compounds of the invention that can be used to treat or prevent cancer include: lats proteins, including human lats proteins, therapeutically or prophylactically effective lats analogs and fragments, and nucleic acids encoding the lats proteins, analogs and fragments.

In a preferred embodiment, the invention provides therapeutic and prophylactic methods for the treatment or prevention of cancer that has been shown to be or may be

refractory to chemotherapy or radiation therapy treatments or treatments based on tumor suppressor genes other than lats.

Also included in the invention are methods of screening lats proteins and lats derivatives and analogs for activity in treating cancer that has been shown to be or may be refractory to chemotherapy or radiation therapy. Additionally, the invention provides methods of screening lats proteins, lats derivatives and fragments, anti-lats antibodies, lats antisense nucleic acids, lats antagonists and inhibitors, and lats-cdc2 complexes for activity in modulating the activity of cdc2.

### Therapeutic Uses

The invention provides for treatment or prevention of cancers refractory to chemotherapy or radiation therapy by administration of a therapeutic compound (termed herein "Therapeutic"). The invention also provides for treatment or prevention of diseases or disorders that can be treated by modulation of cdc2 activity by administration of a Therapeutic of the invention. Such "Therapeutics" include lats proteins and therapeutically or prophylactically effective analogs and fragments thereof; lats-cdc2 complexes; antibodies thereto; nucleic acids encoding the lats proteins, analogs, or fragments, and lats-cdc2 complexes; lats antisense nucleic acids, and lats agonists and antagonists.

In specific embodiments, the therapeutic is a lats protein or lats-cdc2 complex containing a lats protein that is phosphorylated, particularly a lats protein phosphorylated on a serine or threonine residue within 20 residues upstream of an Ala-Pro-Glu consensus in subdomain eight of a lats kinase domain, *e.g.*, a serine corresponding to serine 909 of human lats, as depicted in Figure 12 (SEQ ID NO:2). In another embodiment, the therapeutic is a lats derivative or lats-cdc2 complex containing a lats derivative, in which derivative a serine or threonine residue within 20 residues upstream of an Ala-Pro-Glu consensus in subdomain eight of a lats kinase domain is substituted with a glutamate or aspartate residue, preferably the serine corresponding to serine 909 of human lats is replaced with a glutamate residue. In a further embodiment, the therapeutic is a fragment of a lats protein or a lats-cdc2 complex containing a fragment of a lats protein comprising or consisting of the amino acid sequence of a lats protein corresponding to amino acids 15-585 of human lats, as depicted in Figure 12 (SEQ ID NO:2).

Generally, administration of products of a species origin or species reactivity (in the case of antibodies) that is the same species as that of the subject is preferred. Thus, in a preferred embodiment, a human lats protein, derivative, or fragment, or nucleic acid, or an antibody to a human lats protein, is therapeutically or prophylactically administered to a human patient.

### Treatment and Prevention of Cancers Refractory to Chemotherapy or Radiotherapy

Cancers, including neoplasms, tumors, metastases, or any disorder characterized by uncontrolled cell growth, that have been shown to be refractory to a chemotherapy or

radiation therapy can be treated or prevented by administration of a Therapeutic of the invention that promotes (*i.e.*, increases or supplies) lats function.

5 Examples of such a Therapeutic include lats proteins, derivatives or fragments that are functionally active, particularly have a lats functional activity of inhibiting cell overproliferation (*e.g.*, as demonstrated in *in vitro* assays or in an animal model), and nucleic acids encoding a lats protein or a functionally active derivative or analog thereof (*e.g.*, for use in gene therapy). Other Therapeutics that can be used, *e.g.*, lats agonists, can be identified using *in vitro* assays or animal models, examples of which are described in Examples section.

10 That a cancer is refractory to chemotherapy or radiation therapy means that at least some significant portion of the cancer cells are not killed or their cell division arrested by the particular chemotherapeutic agent or combination of chemotherapeutic agents or the level of radiation employed in a therapeutic protocol. The determination of whether the cancer cells are refractory to the chemotherapy or the radiation therapy can be made either *in vivo* or *in vitro* by any method known in the art for assaying the effectiveness of treatment on cancer cells.

15 In various embodiments of the invention, cancer that is refractory to radiation therapy, chemotherapy or combination chemotherapy, or combination of radiotherapy and chemotherapy, is treated or prevented by administration of a Therapeutic of the invention. In a preferred embodiment, cancer that is refractory to treatment with a chemotherapeutic agent that is cell cycle specific, or said cancer is refractory to treatment with a chemotherapeutic agent that kills or arrests the cells in the S phase of the cell cycle, or said cancer is refractory to treatment with a chemotherapeutic agent that kills or arrests cells during the M phase of the cell cycle is treated using a Therapeutic of the invention. The Therapeutic of the invention can be administered along with radiation therapy and/or one or a combination of chemotherapeutic agents, or as an alternative to other forms of therapy.

25 The chemotherapy or radiation therapy administered concurrently with or subsequent to the administration of the therapeutic of the invention can be administered by any method known in the art. The chemotherapeutic agents are preferably administered in a series of sessions, any one or a combination of the chemotherapeutic agents listed above can be administered. With respect to radiation therapy, any radiation therapy protocol can be used depending upon the type of cancer to be treated. For example, but not by way of limitation, x-ray radiation can be administered; in particular, high-energy megavoltage (radiation of greater than 1 MeV energy) can be used for deep tumors, and electron beam and orthovoltage x-ray radiation can be used for skin cancers. Gamma ray emitting radioisotopes, such as radioactive isotopes of radium, cobalt and other elements may also be administered to expose tissues to radiation.

## Malignancies

Malignancies and related disorders that may become refractory to chemotherapy and/or radiation therapy and that can be treated or prevented by administration of a Therapeutic that promotes lats function include blood-related cancers and solid tumors (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

## Treatment and Prevention of Diseases by Modulation of Cdc2 Activity

In a specific embodiment of the invention, diseases and disorders associated with aberrant levels of cdc2 activity, *e.g.*, aberrantly high or aberrantly low levels of cdc2 protein or activity, can be treated or prevented by administration of a therapeutic of the invention able to modulate the activity of cdc2. In particular, those diseases and disorders associated with an aberrantly high cdc2 activity are treated or prevented by administration of a Therapeutic that promotes lats activity. Alternatively, those diseases and disorders associated with an aberrantly low cdc2 activity are treated or prevented by administration of a Therapeutic that inhibits lats activity, *e.g.*, lats derivatives and analogs that inhibit or antagonize lats activity, anti-lats antibodies, lats antisense nucleic acids, lats inhibitors and antagonists, antibodies that specifically recognize a lats-cdc2 complex, etc.

Because cdc2 promotes cell division, diseases and disorders that may be associated with an increased level of cdc2 activity, as compared with the levels of cdc2 in a subject not afflicted with such a disease or disorder, include diseases and disorders associated with increased cell proliferation, such as malignancies. Diseases and disorders that may be associated with a decreased level of cdc2 activity include diseases and disorders associated with decreased cell proliferation.

## Premalignant Conditions

The Therapeutics of the invention that reduce cdc2 activity can be administered to treat premalignant conditions and to prevent progression to a neoplastic or malignant state. Such prophylactic or therapeutic use is indicated in conditions known or suspected of preceding progression to neoplasia or cancer, in particular, where non-neoplastic cell growth consisting of hyperplasia, metaplasia, or most particularly, dysplasia has occurred.

Alternatively or in addition to the presence of abnormal cell growth characterized as hyperplasia, metaplasia, or dysplasia, the presence of one or more characteristics of a transformed phenotype, or of a malignant phenotype, displayed *in vivo* or displayed *in vitro* by a cell sample from a patient, can indicate the desirability of prophylactic/therapeutic administration of a Therapeutic that inhibits cdc2 activity. Some characteristics of a transformed phenotype include morphology changes, looser substratum attachment, loss of contact inhibition, loss of anchorage dependence, protease release, increased sugar transport, decreased serum requirement, expression of fetal antigens, disappearance of the 250,000 dalton cell surface protein, etc.

### Treatment and Prevention of Disorders in Which Cell Proliferation Is Desired

Diseases and disorders involving a deficiency in cell proliferation (growth) or in which cell proliferation is otherwise desirable for treatment or prevention, are treated or prevented by administration of a Therapeutic that antagonizes (inhibits) *lats* function (in particular, *lats*-mediated inhibition of cell proliferation and/or *lats* binding to *cdc2*).

5 Therapeutics that can be used include anti-*lats* antibodies (and fragments and derivatives thereof containing the binding region thereof), *lats* derivatives or fragments that are dominant-negative kinases, *lats* antisense nucleic acids, and *lats* nucleic acids that are dysfunctional (*e.g.*, due to a heterologous (non-*lats* sequence) insertion within the *lats* coding sequence) that are used to "knockout" endogenous *lats* function by homologous recombination (see, *e.g.*, Capecchi, 1989, Science 244:1288-1292), as described herein. 10 Other Therapeutics that inhibit *lats* function can be identified by use of known convenient *in vitro* assays, *e.g.*, based on their ability to inhibit binding of *lats* to another protein (*e.g.*, *cdc2*), or inhibit any known *lats* function, as preferably assayed *in vitro* or in cell culture. Methods for screening for compounds that prevent or reduce *lats* binding to *cdc2* are 15 described herein. Preferably, suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific Therapeutic (*i.e.*, its ability to promote *cdc2* activity or increase *cdc2* levels) and whether its administration is indicated for treatment of the affected tissue.

Diseases and disorders involving a deficiency in cell proliferation or in which cell proliferation is desired for treatment or prevention, and that can be treated or prevented by promoting *cdc2* function, include degenerative disorders, growth deficiencies, 20 hypoproliferative disorders, physical trauma, lesions, and wounds; for example, to promote wound healing, or to promote regeneration in degenerated, lesioned or injured tissues, etc.

### Gene Therapy

Gene therapy refers to therapy performed by the administration of a nucleic acid to a subject. In this embodiment of the invention, the nucleic acid produces its encoded protein 25 that mediates a therapeutic effect by promoting *lats* function.

Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

For general reviews of the methods of gene therapy, see Goldspiel et al., 1993, Clinical Pharmacy 12:488-505; Wu and Wu, 1991, Biotherapy 3:87-95; Tolstoshev, 1993, 30 Ann. Rev. Pharmacol. Toxicol. 32:573-596; Mulligan, 1993, Science 260:926-932; and Morgan and Anderson, 1993, Ann. Rev. Biochem. 62:191-217; May, 1993, TIBTECH 11(5):155-215. Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), 1993, Current Protocols in Molecular Biology, John Wiley & Sons, NY; and Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY. 35

In a preferred aspect, the Therapeutic comprises a *lats* nucleic acid that is part of an expression vector that expresses a *lats* protein or fragment or chimeric protein thereof in a

suitable host. In particular, such a nucleic acid has a promoter operably linked to the *lats* coding region, said promoter being inducible or constitutive, homologous or heterologous, and, optionally, tissue-specific. In another particular embodiment, a nucleic acid molecule is used in which the *lats* coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the *lats* nucleic acid, as described (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438). In another embodiment, a nucleic acid or combination of nucleic acids containing both a *lats* and a *cdc2* nucleic acid, preferably where each is operably linked to a promoter, is delivered by gene therapy methods.

Delivery of the nucleic acid into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, cells are first transformed with the nucleic acid *in vitro*, then transplanted into the patient. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy.

Numerous techniques are known in the art for the introduction of foreign genes into cells (see e.g., Loeffler and Behr, 1993, Meth. Enzymol. 217:599-618; Cohen et al., 1993, Meth. Enzymol. 217:618-644; Cline, 1985, Pharmac. Ther. 29:69-92).

In an embodiment in which recombinant cells are used in gene therapy, a *lats* nucleic acid or both *lats* and *cdc2* nucleic acids are introduced into the cells such that they are expressible by the cells or their progeny, and the recombinant cells are then

administered *in vivo* for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem and/or progenitor cells which can be isolated and maintained *in vitro* can potentially be used, such as hematopoietic stem cells (HSC), stem cells of epithelial tissues such as the skin and the lining of the gut, embryonic heart muscle cells, liver stem cells (PCT Publication WO 94/08598, dated April 28, 1994), neural stem cells (Stemple and Anderson, 1992, Cell 71:973-985), or epithelial stem cells (ESCs) (Rheinwald, 1980, Meth. Cell Bio. 21A:229; Pittelkow and Scott, 1986, Mayo Clinic Proc. 61:771).

### Antisense Therapy

*Lats* function may be inhibited by use of *lats* antisense nucleic acids. The present invention provides the therapeutic or prophylactic use of nucleic acids of at least six nucleotides and are preferably oligonucleotides (ranging from 6 to about 200 oligonucleotides), that are antisense to a gene or cDNA encoding a *lats* protein, or portions thereof. A *lats* "antisense" nucleic acid as used herein refers to a nucleic acid capable of hybridizing to a portion of a *lats* nucleic acid (preferably mRNA) by virtue of some sequence complementarity. The antisense nucleic acid may be complementary to a coding and/or noncoding region of a *lats* mRNA. In specific aspects, the oligonucleotide is at least 10 nucleotides, at least 15 nucleotides, at least 100 nucleotides, or at least 200 nucleotides.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at any position (examples of such modifications can be found in: Bailey, Ullmann's Encyclopedia of Industrial Chemistry (1998), 6th ed. Wiley and Sons). Such antisense nucleic acids have utility as Therapeutics that inhibit *lats* function or activity, and can be used in the treatment or prevention of disorders characterized by an aberrantly low *cdc2* level or activity.

The *lats* antisense nucleic acids can be directly administered to a cell, or can be produced intracellularly by transcription of exogenous, introduced sequences. Alternatively, *lats* antisense nucleic acids are produced intracellularly by transcription from an exogenous sequence. For example, a vector can be introduced *in vivo* such that it is taken up by a cell, within which cell the vector or a portion thereof is transcribed, producing an antisense nucleic acid (RNA) of the invention. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art.

The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a *lats* gene, preferably a human *lats* gene. However, absolute complementarity, although preferred, is not required.

Pharmaceutical compositions of the invention, comprising an effective amount of a *lats* antisense nucleic acid in a pharmaceutically acceptable carrier can be administered to a patient having a disease or disorder which is characterized by aberrantly low *cdc2* activity.

The amount of *lats* antisense nucleic acid that will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. Where possible, it is desirable to determine the antisense cytotoxicity *in vitro*, and then in useful animal model systems prior to testing and use in humans.

In a specific embodiment, pharmaceutical compositions comprising *lats* antisense nucleic acids are administered via liposomes, microparticles, or microcapsules. In various embodiments of the invention, it may be useful to use such compositions to achieve sustained release of the *lats* antisense nucleic acids. In a specific embodiment, it may be desirable to utilize liposomes targeted via antibodies to specific identifiable central nervous system cell types (Leonetti et al., 1990, Proc. Natl. Acad. Sci. U.S.A. 87: 2448-2451; Renneisen et al., 1990, J. Biol. Chem. 265: 16337-16342).

#### **Lats Proteins, Derivatives, Fragments and Lats-cdc2 Complexes**

The *lats* proteins and nucleic acids, and *lats* derivatives and fragments can be produced by any method known in the art.

For recombinant expression of *lats* proteins, and *lats* derivatives and fragments, the nucleic acid containing all or a portion of the nucleotide sequence encoding the protein can

be inserted into an appropriate expression vector, *i.e.*, a vector that contains the necessary elements for the transcription and translation of the inserted protein coding sequence. In a preferred embodiment, the regulatory elements (*e.g.*, promoter) are heterologous (*i.e.*, not the native gene promoter). Promoters which may be used include the SV40 early promoter (Bernoist and Chambon, 1981, Nature 290: 304-310), and the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980, Cell 22: 787-797), among others.

A variety of host-vector systems may be utilized to express the protein coding sequence. These include mammalian cell systems infected with virus (*e.g.*, vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (*e.g.* baculovirus); microorganisms such as yeast containing yeast vectors; or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA.

Once a lats protein, or derivative or fragment, has been recombinantly expressed, it may be isolated and purified by standard methods including chromatography (*e.g.*, ion exchange, affinity, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins. A lats protein may also be purified by any standard purification method from natural sources.

Alternatively, a lats protein, analog or derivative can be synthesized by standard chemical methods known in the art (*e.g.*, see Hunkapiller et al., 1984, Nature 310:105-111).

The Therapeutics of the invention also include derivatives and fragments related to lats. In particular embodiments, the derivative or fragment is functionally active, *i.e.*, capable of exhibiting one or more functional activities associated with a full-length, wild-type lats protein, *e.g.*, able to inhibit cell proliferation in *in vitro* and/or *in vivo* assays. Additionally, derivatives or fragments that inhibit lats activity, *e.g.*, promote cell proliferation, may also have a use in the methods of the invention. Derivatives or analogs of lats can be tested for the desired activity by procedures known in the art.

In specific embodiments of the invention, the Therapeutic is a lats protein that is phosphorylated, preferably that is phosphorylated on a serine or threonine residue within 20 amino acids upstream of an Ala-Pro-Glu consensus sequence in subdomain eight of a lats kinase domain, more preferably that is phosphorylated on a serine residue corresponding to serine 909 of human lats, as depicted in Figure 12 (SEQ ID NO:2). In another specific embodiment, the therapeutic is a lats derivative in which a serine or threonine residue within 20 residues upstream of an Ala-Pro-Glu consensus in subdomain eight of a lats kinase domain is substituted with a glutamate or aspartate residue, preferably, in which a serine residue corresponding to serine 909 of human lats is replaced with a glutamate residue. In another specific embodiment, the therapeutic is a fragment of a lats protein comprising or consisting of the amino acid sequence corresponding to amino acids 15 to 585 of human lats, as depicted in Figure 12 (SEQ ID NO:2).

In particular, lats derivatives can be made by altering lats sequences by substitutions, additions or deletions that provide for functionally equivalent molecules. Due



to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence as a *lats* gene may be used in the practice of the present invention. These include nucleotide sequences comprising all or portions of *lats* genes which are altered by the substitution of different codons that encode a functionally equivalent amino acid residue within the sequence, thus producing a silent change.

5 Likewise, the *lats* derivatives of the invention include those containing, as a primary amino acid sequence, all or part of the amino acid sequence of a *lats* protein including altered sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence resulting in a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity  
10 which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and  
15 histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

In a specific embodiment of the invention, proteins consisting of or comprising a fragment of a *lats* protein consisting of at least 10 (continuous) amino acids of the *lats* protein is provided. In other embodiments, the fragment consists of at least 20 or 50 amino  
20 acids of the *lats* protein. In specific embodiments, such fragments are not larger than 35, 100 or 200 amino acids. In a specific embodiment, the fragment of a *lats* protein is from the N-terminal portion of the protein, preferably including all or a portion of the amino acids corresponding to amino acids 15-585 of human *lats*. Derivatives or fragments of *lats* include but are not limited to those molecules comprising regions that are substantially  
25 homologous to *lats* or fragments thereof (*e.g.*, in various embodiments, at least 60% or 70% or 80% or 90% or 95% identity over an amino acid sequence of identical size with no insertions or deletions considered, or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, *e.g.*, the blastp program) or whose encoding nucleic acid is capable of hybridizing to the inverse complement (the inverse complement of a nucleic acid strand has the complementary  
30 sequence running in reverse orientation to the strand so that the inverse complement would hybridize without mismatches to the nucleic acid strand; thus, for example, where the coding strand is hybridizable to a nucleic acid with no mismatches between the coding strand and the hybridizable strand, then the inverse complement of the hybridizable strand is identical to the coding strand) of a coding *lats* sequence, under high stringency, moderately stringency, or low stringency conditions, as discussed *infra*.

35 The *lats* derivatives and fragments of the invention can be produced by various methods known in the art. The manipulations which result in their production can occur at

the gene or protein level. For example, the cloned *lats* gene sequence can be modified by any of numerous strategies known in the art (Sambrook et al., 1990, Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

5 Additionally, the *lats*-encoding nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification. Any technique for mutagenesis known in the art can be used, including chemical mutagenesis, *in vitro* site-directed mutagenesis (Hutchinson, C., et al., 1978, J. Biol. Chem 253:6551), use of TAB® linkers  
10 (Pharmacia), etc.

Manipulations of the *lats* sequence may also be made at the protein level. Included within the scope of the invention are *lats* protein fragments or other derivatives which are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, etc. Any of numerous chemical modifications may be carried out by  
15 known techniques, including specific chemical cleavage by cyanogen bromide, trypsin, oxidation, reduction; etc.

In addition, analogs and fragments of *lats* can be chemically synthesized. For example, a peptide corresponding to a portion of a *lats* protein which comprises the desired domain, or which mediates the desired activity *in vitro*, can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino  
20 acid analogs can be introduced as a substitution or addition into the *lats* sequence.

In a specific embodiment, the *lats* derivative is a chimeric, or fusion, protein comprising a *lats* protein or fragment thereof (preferably consisting of at least a domain or motif of the *lats* protein, or at least 15, preferably 20, amino acids of the *lats* protein) joined at its amino- or carboxy-terminus via a peptide bond to an amino acid sequence of a  
25 different protein. In one embodiment, such a chimeric protein is produced by recombinant expression of a nucleic acid encoding the protein (comprising a *lats*-coding sequence joined in-frame to a coding sequence for a different protein). Such a chimeric product can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other by methods known in the art, in the proper coding frame, and expressing the chimeric product by methods commonly known in the art. Alternatively,  
30 such a chimeric product may be made by protein synthetic techniques, *e.g.*, by use of a peptide synthesizer. Chimeric genes comprising portions of *lats* fused to any heterologous protein-encoding sequences may be constructed. A specific embodiment relates to a chimeric protein comprising a fragment of *lats* of at least six amino acids. In another specific embodiment, the *lats* derivative is a chimeric protein comprising a fragment of *lats*  
35 corresponding to amino acids 15-585 of human *lats*.

In another specific embodiment, the *lats* derivative is a molecule comprising a region of homology with a *lats* protein. By way of example, in various embodiments, a first

protein region can be considered "homologous" to a second protein region when the amino acid sequence of the first region is at least 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, or 95% identical, when compared to any sequence in the second region of an equal number of amino acids as the number contained in the first region with no insertions or deletions considered, or when compared to an aligned sequence of the second region that has been aligned by a computer homology program known in the art. For example, a molecule can comprise one or more regions homologous to a lats domain or a portion thereof.

### Derivatives of Lats Containing One or More Domains of the Protein

In specific embodiments, the methods of the invention use lats derivatives and fragments that comprise, or alternatively consist of, one or more domains of a lats protein, including but not limited to a lats C-terminal domain 3 (LCD3), lats C-terminal domain 2 (LCD2), lats C-terminal domain 1 (LCD1), kinase domain, kinase subdomains, lats flanking domain (LFD) (amino-terminal to the kinase domain), lats split domain 1 (LSD1), lats split domain 2 (LSD2), SH3-binding domain, and opa repeat domain, functional (*e.g.*, binding) fragments of any of the foregoing, or any combination of the foregoing.

In human lats (h-lats), m-lats, m-lats2, and *Drosophila* lats, the LCD3 domain is the last three amino acids of the protein, which are Val-Tyr-Val in all four proteins. For human lats and *Drosophila* lats, the LCD2 domain is amino acid residues 1077-1086 and 1075-1084, respectively (all amino acid residues provided in this paragraph are for the human and *Drosophila* lats amino acid sequences depicted in Figures 12 and 14, respectively (SEQ ID NOS:2 and 8, respectively)); the LCD1 domain is amino acid residues 1032-1043 and 1035-1047, respectively; the kinase domain is amino acid residues 703-1014 and 711-1018, respectively; the LFD domain is amino acid residues 607-702 and 612-710 respectively; and the putative SH3-binding domain is amino acids 247-268 and 196-217, respectively. For the lats split domains in *Drosophila*, the LSD1 is amino acid residues 365-392 and the LSD2 is amino acids 536-544. In human lats, the LSD1 and LSD2 domains are split into anterior and posterior portions such that the LSD1 is amino acid residues 328-334 and 498-518 and LSD2 is amino acid residues 28-31 and 555-559.

In particular, the Therapeutics of the invention include molecules comprising specific fragments of lats that are those fragments in the respective lats protein most homologous to specific fragments of a human or mouse lats protein.

In a specific embodiment, a lats protein, derivative or fragment is provided that has a kinase domain and has a phosphorylated or dephosphorylated serine situated within 20 residues upstream of an Ala-Pro-Glu consensus in subdomain eight of its kinase domain, or in which the serine situated within 20 residues upstream of that consensus has been deleted or substituted by another amino acid. In various specific embodiments, the invention provides various phosphorylated and dephosphorylated forms of the lats protein, derivative, or fragment that are active or inactive kinase forms. Both phosphorylation and dephosphorylation of lats at different residues could potentially activate or inactivate lats.

Phosphorylation can be carried out by any methods known in the art, *e.g.*, by use of a kinase. Dephosphorylation can be carried out by use of any methods known in the art, *e.g.*, by use of a phosphatase.

Another specific embodiment relates to a derivative or fragment of a lats protein that is a dominant-active protein kinase. Such a derivative or analog comprises a lats kinase domain that has been mutated so as to be dominantly active (exhibit constitutively active kinase activity). It is known that acidic residues such as Glu and Asp sometimes mimic a phosphorylated residue, and changing the phosphorylatable Ser or Thr residue in subdomain eight into a Glu or Asp residue has been previously used to produce constitutively active kinases (Mansour et al., 1994, Science 265:966-970). Thus, changing a serine or threonine residue situated within 20 residues upstream of an Ala-Pro-Glu consensus in subdomain eight of a lats kinase domain into another residue (*e.g.*, Glu, Asp) may be used to make a dominant-active lats protein kinase. For example, changing Ser914 in *Drosophila* lats, or changing Ser909 in human lats, into a Glu residue could produce a dominant active lats kinase.

Another specific embodiment relates to a derivative or fragment of lats that is a dominant-negative protein kinase. Protein kinases can be mutated into dominant negative forms. Expression of a dominant negative protein kinase can suppress the activity of the wild-type form of the same kinase. Dominant negative forms of protein kinases are often obtained by expressing an inactive form of a kinase (Milarski and Saltiel, 1994, J. Biol. Chem. 269(33):21239-21243) or by expressing a noncatalytic domain of a kinase (Lu and Means, 1994, EMBO J. 12:2103-2113; Yarden et al., 1992, EMBO J. 11:2159-2166). Thus, a lats dominant-negative kinase can be obtained by mutating the kinase domain so as to be inactive (*e.g.*, by deletion and/or point mutation). For example, a lats derivative that is a dominant-negative kinase is a lats protein that lacks a kinase domain but comprises one or more of the other domains of the lats protein; *e.g.*, a lats protein derivative truncated at about the beginning of the kinase domain (*i.e.*, a lats fragment containing only sequences amino-terminal to the kinase domain). As another example, a lats derivative that is a dominant-negative kinase is a lats protein in which one of the residues conserved among serine/threonine kinases (see Hanks et al., 1988, Science 241:42-52) is mutated (deleted or substituted by a different residue).

In another specific embodiment, a molecule is provided that comprises one or more domains (or functional portion thereof) of a lats protein but that also lacks one or more domains (or functional portion thereof) of a lats protein. For example, such a protein may lack all or a portion of the kinase domain, but retain at least the SH3-binding domain of a lats protein. In another embodiment, a molecule is provided that comprises one or more domains (or functional portion thereof) of a lats protein, and that has one or more mutant (*e.g.*, due to deletion or point mutation(s)) domains of a lats protein (*e.g.*, such that the mutant domain has decreased function). The kinase domain may be mutant so as to have reduced, absent, or increased kinase activity.

### Lats-cdc2 Complexes

The invention provides lats-cdc2 complexes. In a preferred embodiment, the lats-cdc2 complexes are complexes of human proteins. As used herein, fragment or derivative of a lats-cdc2 complex includes complexes where one or both members of the complex are fragments or derivatives of the wild-type lats or cdc2 protein. Such derivatives and fragments can be generated as described for lats derivatives and fragments above. Preferably, the lats-cdc2 complexes in which one or both members of the complex are a fragment or derivative of the wild type protein are functionally active lats-cdc2 complexes.

In particular aspects, the native proteins, derivatives or analogs of lats and/or cdc2 are of animals, *e.g.* mouse, rat, pig, cow, dog, monkey, human, fly, frog, or of plants.

“Functionally active lats-cdc2 complex” as used herein refers to that material displaying one or more known functional attributes of a complex of full length lats with a full length cdc2, including but not exclusive to control of cell cycle progression, cell proliferation, etc.

In specific embodiments, the lats-cdc2 complex contains a lats protein that is phosphorylated, preferably that is phosphorylated on a serine or threonine residue within 20 amino acids upstream of an Ala-Pro-Glu consensus subdomain eight of a lats kinase domain, more preferably that is phosphorylated on a serine residue corresponding to serine 909 of human lats, as depicted in Figure 12 (SEQ ID NO:2). In another specific embodiment, the lats-cdc2 complex contains a lats derivative in which a serine or threonine residue within 20 residues upstream of an Ala-Pro-Glu consensus subdomain eight of a lats kinase domain is substituted with a glutamate or aspartate residue, preferably, in which a serine residue corresponding to serine 909 of human lats is replaced with a glutamate residue. In another specific embodiment, the therapeutic is a fragment of a lats protein comprising or consisting of the amino acid sequence corresponding to amino acids 15 to 585 of human lats, as depicted in Figure 12 (SEQ ID NO:2).

Methods are presented for screening lats-cdc2 complexes, as well as derivatives and fragments of the lats-cdc2 complexes for the ability to alter lats and/or cdc2 activity, *e.g.*, to alter cell proliferation. For example, such derivatives or fragments which have the desired immunogenicity or antigenicity can be used in immunoassays, for immunization, for inhibition of lats-cdc2 complex activity, etc. Derivatives or fragments that retain, or alternatively lack or inhibit, a property of interest (*e.g.*, participation in a lats-cdc2 complex) can be used as inducers, or inhibitors, respectively, of such a property and its physiological correlates. A specific embodiment relates to a lats-cdc2 complex of a fragment of lats and/or a fragment of cdc2 that can be bound by an anti-lats and/or anti-cdc2 antibody or antibody specific for a lats-cdc2 complex when such a fragment is included within a lats-cdc2 complex.

The lats-cdc2 complexes can be obtained by any method known in the art. The cdc2 nucleotide and amino acid sequence is available from GenBank, accession no. Y00272 (see also, Lee and Nurse, 1987, Nature 327:31-35). The lats-cdc2 complexes can be obtained, for example, by expressing an entire lats coding sequence and a cdc2 coding sequence in the

same cell, either under the control of the same promoter or two separate promoters. In yet another embodiment, a derivative, fragment or homolog of lats and/or a derivative, fragment or homolog of cdc2 are recombinantly expressed. Preferably, the derivative, fragment or homolog of lats and/or the cdc2 protein form a complex with a binding partner identified by a binding assay, such as co-immunoprecipitation with an anti-lats or anti-cdc2 antibody, or interaction in a yeast two-hybrid assay (Fields and Song, 1989, Nature 340:245-246; and Finley and Brent, in DNA Cloning 2, Rickwood and Hames, eds (Oxford University Press, Oxford, 1995)).

In a specific embodiment, fusion or chimeric proteins are provided that contain the domains of a lats protein, or, in a specific embodiment, the amino acid sequence corresponding to amino acids 15 to 585 of human lats, and a cdc2 protein that directly form a lats-cdc2 complex and, optionally, a heterofunctional reagent, such as a peptide linker, linking the two domains, where such a heterofunctional reagent, such as a reagent or linker promotes the interaction of the lats and cdc2 binding domains. These fusion proteins may be particularly useful where the stability of the interaction is desirable (due to the formation of the complex as an intramolecular reaction), for example in production of antibodies specific to the lats-cdc2 complex.

#### **Generation of Antibodies to Lats Proteins and Lats-cdc2 Complexes and Derivatives**

LATS proteins, including functional derivatives and fragments thereof (*e.g.* a LATS protein encoded by a sequence of any one of SEQ ID NOs:2, 4, 6, or 8, or a subsequence thereof) may be used as an immunogen to generate monoclonal or polyclonal antibodies and antibody fragments or derivatives (*e.g.*, chimeric, single chain, Fab fragments, etc.). For example, antibodies to a particular domain of a lats protein may be desired. In a specific embodiment, fragments of a lats protein identified as hydrophilic are used as immunogens for antibody production using art-known methods. Various known methods for antibody production can be used including cell culture of hybridomas; production of monoclonal antibodies in germ-free animals (PCT/US90/02545); the use of human hybridomas (Cole *et al.*, Proc. Natl. Acad. Sci. U.S.A. (1983) 80:2026-2030; Cole *et al.*, in Monoclonal Antibodies and Cancer Therapy (1985) Alan R. Liss, pp. 77-96), and production of humanized antibodies (Jones *et al.*, Nature (1986) 321:522-525; US Pat. No. 5,530,101).

#### **Diagnostic, Prognostic, and Screening Uses of Lats-cdc2 Complexes**

Lats-cdc2 complexes may be markers of specific disease states involving disruption of physiological processes, such as cell cycle progression and cell proliferation, and pathological processes, such as hyperproliferative disorders, including tumorigenesis and tumor progression, and hypoproliferative disorders, and thus have diagnostic utility.

Detecting levels of lats-cdc2 complexes, or individual lats and cdc2 proteins or the mRNA encoding lats and cdc2 may be used in diagnosis or prognosis, to follow the course of disease states, or to follow therapeutic response, etc.

Lats-cdc2 complexes, lats and cdc2 proteins, and derivatives, and sub-sequences thereof, *lats* and/or *cdc2* nucleic acids (and sequences complementary thereto), and anti-lats-cdc2 complex antibodies and combinations of antibodies directed against lats and cdc2 have uses in diagnostics. Such molecules can be used in assays, such as immunoassays, to detect, prognose, diagnose, or monitor various conditions, diseases, and disorders characterized by aberrant levels of lats-cdc2 complexes or monitor the treatment thereof.

In particular, such an immunoassay is carried out by a method comprising contacting a sample derived from a patient with an anti-lats-cdc2 complex antibody under conditions such that immunospecific binding can occur, and detecting or measuring the amount of any immunospecific binding by the antibody. In a specific aspect, such binding of antibody, in tissue sections, can be used to detect aberrant lats-cdc2 complex localization or aberrant (*e.g.*, high, low or absent) levels of lats-cdc2 complex. In a specific embodiment, an antibody to a lats-cdc2 complex can be used to assay in a patient tissue or serum sample for the presence of a lats-cdc2 complex where an aberrant level of lats-cdc2 complex is an indication of a diseased condition. By "aberrant levels" is meant an increased or decreased level relative to that present, or a standard level representing that present, in an analogous sample from a portion of the body or from a subject not having the disorder.

The immunoassays which can be used include competitive and non-competitive assay systems using techniques such as Western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, etc.

Nucleic acids encoding lats and cdc2 proteins and related nucleotide sequences and sub-sequences, including complementary sequences, can also be used in hybridization assays. The *lats* and *cdc2* nucleotide sequences, or sub-sequences thereof comprising about at least 8 nucleotides, can be used as hybridization probes. Hybridization assays can be used to detect, prognose, diagnose, or monitor conditions, disorders, or disease states associated with aberrant levels of the mRNAs encoding the components of a lats-cdc2 complex. In particular, such a hybridization assay is carried out by a method comprising contacting a sample containing nucleic acid with a nucleic acid probe capable of hybridizing to *lats* and *cdc2* DNAs or RNAs, under conditions such that hybridization can occur, and detecting or measuring any resulting hybridization.

By way of example, levels of lats-cdc2 complexes and lats and cdc2 proteins can be detected by immunoassay, levels of lats and cdc2 mRNA can be detected by hybridization assays (*e.g.*, Northern blots, dot blots), binding of lats to cdc2 can be done by binding assays commonly known in the art, translocations and point mutations in *lats* and/or *cdc2* can be detected by Southern blotting, RFLP analysis, PCR using primers that preferably

generate a fragment spanning at least most of the *lats* and/or *cdc2* gene, sequencing of the *lats* and/or *cdc2* genomic DNA or cDNA obtained from the patient, etc.

Also embodied are methods to detect a lats-cdc2 complex in cell culture models that express a lats-cdc2 complex, or derivatives thereof, for the purpose of characterizing or preparing the lats-cdc2 complex for harvest. This embodiment includes cell sorting of prokaryotes such as but not restricted to, bacteria (Davey and Kell, 1996, Microbiol. Rev. 60:641-696), primary cultures and tissue specimens from eukaryotes, including mammalian species such as human (Steele et al., 1996, Clin. Obstet. Gynecol 39:801-813), and continuous cell cultures (Orfao and Ruiz-Arguelles, 1996, Clin. Biochem. 29:5-9).

Kits for diagnostic use are also provided, that comprise in one or more containers an anti-lats-cdc2 complex antibody and, optionally, a labeled binding partner to the antibody. Alternatively, the anti-lats-cdc2 complex antibody can be labeled (with a detectable marker, e.g., a chemiluminescent, enzymatic, fluorescent, or radioactive moiety). A kit is also provided that comprises in one or more containers a nucleic acid probe or probes capable of hybridizing to *lats* and *cdc2* mRNAs. In a specific embodiment, a kit can comprise in one or more containers a pair of primers (e.g., each in the size range of 6-30 nucleotides) that are capable of priming amplification [e.g., by polymerase chain reaction (see e.g., Innis et al., 1990, PCR Protocols, Academic Press, Inc., San Diego, CA), ligase chain reaction (see EP 320,308) use of  $\beta$ -replicase, cyclic probe reaction, or other methods known in the art], under appropriate reaction conditions of at least a portion of a *lats* nucleic acid and a *cdc2* nucleic acid. A kit can optionally further comprise in a container a predetermined amount of a purified lats-cdc2 complex, lats and cdc2 proteins or nucleic acids thereof, e.g., for use as a standard or control.

### Demonstration of Therapeutic Utility

The Therapeutics of the invention are preferably tested *in vitro*, and then *in vivo*, for the desired therapeutic or prophylactic activity, prior to use in humans. For example, *in vitro* assays which can be used to determine whether administration of a specific Therapeutic is indicated, include *in vitro* cell culture assays in which a patient tissue sample is grown in culture, and exposed to or otherwise administered a Therapeutic, and the effect of such Therapeutic upon the tissue sample is observed. A lower level of proliferation or survival of the contacted cells indicates that the Therapeutic is effective to treat the condition in the patient. Alternatively, instead of culturing cells from a patient, Therapeutics may be screened using cells of a tumor or malignant cell line. Many assays standard in the art can be used to assess such survival and/or growth; for example, cell proliferation can be assayed by measuring  $^3\text{H}$ -thymidine incorporation, by direct cell count, by detecting changes in transcriptional activity of known genes such as proto-oncogenes (e.g., *fos*, *myc*) or cell cycle markers; cell viability can be assessed by trypan blue staining, differentiation can be assessed visually based on changes in morphology, etc.



In another embodiment of the invention, a Therapeutic of the invention is screened for activity to modulate (*e.g.*, promote, inhibit or antagonize) *cdc2* levels and/or activity. The levels of *cdc2* protein and mRNA and *cdc2* activity can be determined by any method well known in the art. For example, *cdc2* protein can be quantitated by known immunodiagnostic methods such as western blotting immunoprecipitation using any antibody against *cdc2* (for example, anti-*cdc2* antibodies are commercially available from Santa Cruz Inc.) *Cdc2* mRNA can be quantitated by methods that are well known and routine in the art, for example by northern analysis, RNase protection, the polymerase chain reaction in connection with the reverse transcription, etc. *Cdc2* activity can also be assayed by any method known in the art, for example, by the histone-H1 kinase assay.

Compounds for use in therapy can be tested in suitable animal model systems prior to testing in humans, including but not limited to in rats, mice, chicken, cows, monkeys, rabbits, etc.

In a preferred embodiment, *lats* knock-out mice, *e.g.*, as described in the Examples, are used to test therapeutics of the invention for activity to treat or prevent cancers, or to modulate *cdc2* activity.

#### ***Lats* Knock-out Animals**

The invention provides recombinant non-human animals in which one or more *lats* genes have been inactivated, *e.g.*, "knock-out animals". The recombinant non-human animal can be any animal, *e.g.*, mouse, rats, rodents, hamster, sheep, pig, cow, *Drosophila*, *C. elegans*, insects, worms, primates, dogs, etc., and is preferably a mouse. Such an animal can be generated by any method known in the art for disrupting a gene on the chromosome of an animal. *Lats* knock-out animals do not include animals in which one or more *lats* genes have been inactivated by naturally occurring mutations. In a preferred aspect, a *lats* knock-out animal can be produced by promoting homologous recombination between a *lats* gene in its chromosome and an exogenous *lats* gene that has been rendered biologically inactive (preferably by insertion of a heterologous sequence, *e.g.*, an antibiotic resistance gene). Homologous recombination methods for disrupting genes in the mouse genome are described, for example, in Capecchi (1989, Science 244:1288-1292) and Mansour et al. (1988, Nature 336:348-352). A *lats* knock-out mouse may be produced by the method described in the Examples section.

Briefly, all or a portion of a *lats* genomic clone is isolated from genomic DNA from the same species as the knock-out animal. The *lats* genomic clone can be isolated by any method known in the art for isolation of genomic clones (*e.g.*, by probing a genomic library with a probe derived from a *lats* sequence, such as those sequences provided in Figures 12-15, *i.e.*, SEQ ID NOS:1, 3, 5, or 7). Once the genomic clone is isolated, all or a portion of the clone is introduced into a recombinant vector. Preferably, the portion of the clone introduced into the vector that contains at least a portion of an exon of the *lats* gene, *i.e.*, contains a *lats* protein coding sequence. A sequence not homologous to the *lats* sequence,

preferably a positive selectable marker, such as a gene encoding an antibiotic resistance gene, is then introduced into the *lats* gene exon. The selectable marker is preferably operably linked to a promoter, more preferably a constitutive promoter. The non-homologous sequence is introduced anywhere in the *lats* coding sequence that will disrupt *lats* activity, *e.g.*, at a position where point mutations or other mutations have been demonstrated to inactivate *lats* protein function. For example, the non-homologous sequence can be inserted for the coding sequence for the portion of the *lats* protein containing all or a portion of the kinase domain (*e.g.*, the nucleotide sequence coding for at least 50, 100, 150, 200 or 250 amino acids of the kinase domain), the Lats C-terminal domain 1, the Lats C-terminal domain 2, and the Lats C-terminal domain 3, or, more preferably, for the sequence coding for the amino acids corresponding to 756 to 1130 of human *lats* (as depicted in Figure 12 (SEQ ID NO:2) and as indicated in the alignment of human and mouse *lats* in Figure 6A).

The positive selectable marker is preferably a neomycin resistance gene (neo gene) or a hygromycin resistance gene (hygro gene). The promoter may be any promoter known in the art; by way of example the promoter may be the phosphoglycerate kinase (PKG) promoter (Adra et al., 1987, Gene 60:65-74), the PolIII promoter (Soriano et al., 1991. Cell 64:693-701), or the MC1 promoter, which is a synthetic promoter designed for expression in embryo-derived stem cells (Thomas & Capecchi, 1987, Cell 51:503-512). Use of a selectable marker, such as an antibiotic resistance gene, allows for the selection of cells that have incorporated the targeting vector (for example, the expression of the neo gene product confers resistance to G418, and expression of the hygro gene product confers resistance to hygromycin).

In a preferred embodiment, a negative selectable marker for a counterselection step for homologous, as opposed to non-homologous, recombination of the vector is inserted outside of the *lats* genomic clone insert, *e.g.*, as shown in Figure 6B. For example, such a negative selectable marker is the HSV thymidine kinase gene (HSV-tk), the expression of which makes cells sensitive to ganciclovir. The negative selectable marker is preferably under the control of a promoter such as the PGK promoter, the PolIII promoter or the MC1 promoter.

When homologous recombination occurs, the portions of the vector that are homologous to the *lats* gene, as well as the non-homologous insert within the *lats* gene sequences, are incorporated into the *lats* gene in the chromosome, and the remainder of the vector is lost. Thus, since the negative selectable marker is outside the region of homology with the *lats* gene, cells in which homologous recombination has occurred (or their progeny), will not contain the negative selectable marker. For example, if the negative selectable marker is the HSV-tk gene, the cells in which homologous recombination has occurred will not express thymidine kinase and will survive exposure to ganciclovir. This procedure permits the selection of cells in which homologous recombination has occurred, as compared to non-homologous recombination in which it is likely that the negative

selectable marker is also incorporated into the genome along with the *lats* sequences and the positive selectable marker. Thus, cells in which non-homologous recombination has occurred would most likely express thymidine kinase and be sensitive to ganciclovir.

Once the targeting vector is prepared, it is linearized with a restriction enzyme for which there is a unique site in the targeting vector, and the linearized vector is introduced into embryo-derived stem (ES) cells (Gossler et al., 1986, Proc. Natl. Acad. Sci. USA 83:9065-9069) by any method known in the art, for example by electroporation. If the targeting vector includes a positive selectable marker and a negative, counterselectable marker, the ES cells in which homologous recombination has occurred can be selected by incubation in selective media. For example, if the selectable markers are the neo resistance gene and the HSV-tk gene, the cells are exposed to G418 (e.g., approximately 300 µg/ml) and ganciclovir (e.g., approximately 2 µM).

Any technique known in the art for genotyping, for example Southern blot analysis or the polymerase chain reaction, can be used to confirm that the disrupted *lats* sequences have homologously recombined into the *lats* gene in the genome of the ES cells. Because the restriction map of the *lats* genomic clone is known (see Figure 6b) and the sequence of the *lats* coding sequence is known (see Figure 13), the size of a particular restriction fragment or a PCR amplification product generated from DNA from both the disrupted and non-disrupted alleles can be determined. Thus, by assaying for a restriction fragment or PCR product, the size of which differs between the disrupted and non-disrupted *lats* gene, one can determine whether homologous recombination has occurred to disrupt the *lats* gene.

The ES cells with the disrupted *lats* locus can then be introduced into mouse blastocysts by microinjection and then the blastocysts can be implanted into the uteri of pseudopregnant mice using routine techniques. The mice that develop from the implanted blastocysts are chimeric for the disrupted allele. The chimeric male mice can be crossed to female mice, and this cross can be designed such that germline transmission of the allele is linked to transmission of a certain coat color. The germline transmission of the allele can be confirmed by Southern blotting or PCR analysis, as described above, of genomic DNA isolated from tail samples.

### Isolating *Lats* Genes

Clones comprising *lats* nucleotide sequences, particularly *lats* genomic clones, can be isolated by any method known in the art. The nucleotide sequences encoding, and the corresponding amino acid sequences of, human *lats*, mouse *lats*, mouse *lats2* and *Drosophila* *lats* are provided in Figures 12-15, respectively (SEQ ID NOS:1-8, respectively) and bacterial cells containing the plasmid pBS(KS)-h-*lats*, which contains the gene encoding human *lats*, were deposited on March 24, 1995 with the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2201, and assigned Accession No. 69769. *Lats* nucleic acids, either *lats* genomic clones or *lats* specific probes to identify *lats* genomic clones, can be obtained by any method known in the art, e.g., from

the deposited plasmid, by the polymerase chain reaction (PCR) using synthetic primers hybridizable to the 3' and 5' ends of a *lats* nucleotide sequence and/or by cloning from a cDNA or genomic library using an oligonucleotide probe specific for the gene sequence, such as a probe from the *lats* gene insert in plasmid pBS(KS)-h-*lats*. Genomic clones can be identified by probing a genomic DNA library under appropriate hybridization conditions, *e.g.*, high stringency conditions, low stringency conditions or moderate stringency conditions, depending on the relatedness of the probe to the genomic DNA being probed. For example, if the *lats* probe and the genomic DNA are from the same species, then high stringency hybridization conditions may be used; however, if the *lats* probe and the genomic DNA are from different species, then low stringency hybridization conditions may be used. High, low and moderate stringency conditions are all well known in the art.

Procedures for low stringency hybridization are as follows (see also Shilo and Weinberg, 1981, Proc. Natl. Acad. Sci. USA 78:6789-6792): Filters containing DNA are pretreated for 6 hours at 40°C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 µg/ml denatured salmon sperm DNA. Hybridizations are carried out in the same solution with the following modifications: 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 µg/ml salmon sperm DNA, 10% (wt/vol) dextran sulfate, and 5-20 X 10<sup>6</sup> cpm <sup>32</sup>P-labeled probe is used. Filters are incubated in hybridization mixture for 18-20 hours at 40°C, and then washed for 1.5 hours at 55°C in a solution containing 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS. The wash solution is replaced with fresh solution and incubated an additional 1.5 hours at 60°C. Filters are blotted dry and exposed for autoradiography. If necessary, filters are washed for a third time at 65-68°C and reexposed to film.

Procedures for high stringency hybridizations are as follows: Prehybridization of filters containing DNA is carried out for 8 hours to overnight at 65°C in buffer composed of 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 µg/ml denatured salmon sperm DNA. Filters are hybridized for 48 hours at 65°C in prehybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 X 10<sup>6</sup> cpm of <sup>32</sup>P-labeled probe. Washing of filters is done at 37°C for 1 hour in a solution containing 2X SSC, 0.01% PVP, 0.01% Ficoll, and 0.01% BSA. This is followed by a wash in 0.1 X SSC at 50°C for 45 minutes before autoradiography.

Moderate stringency conditions for hybridization are as follows: Filters containing DNA are pretreated for 6 hours at 55°C in a solution containing 6X SSC, 5X Denhardt's solution, 0.5% SDS, and 100 µg/ml denatured salmon sperm DNA. Hybridizations are carried out in the same solution and 5-20 X 10<sup>6</sup> cpm <sup>32</sup>P-labeled probe is used. Filters are incubated in the hybridization mixture for 18-20 hours at 55°C, and then washed twice for 30 minutes at 60°C in a solution containing 1 X SSC and 0.1% SDS.

Any eukaryotic cell potentially can serve as the nucleic acid source for the molecular cloning of the *lats* gene. The nucleic acid sequences encoding *lats* can be isolated from vertebrate, mammalian, human, porcine, bovine, feline, avian, equine, canine, as well as

additional primate sources, insects, etc. The DNA may be obtained by standard procedures known in the art, preferably from cloned genomic DNA (e.g., a DNA "library") from the desired cell (see, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II.). The gene should be molecularly cloned into a suitable vector for propagation of the gene.

In preferred embodiments, the genomic clone used to generate a recombinant, non-human animal by homologous recombination contains at least a portion of the *lats* coding sequence of SEQ ID NO:3; alternatively, the genomic clone contains at least a portion of the *lats* coding sequence of SEQ ID NO:5.

### Methods of Screening Therapeutics Using *Lats* Knock-out Mice

The invention provides methods for screening for compounds useful in the treatment or prevention of cancer or in the treatment and prevention of pituitary diseases and disorders by administration or application of the compound to be tested to a *lats* knock-out animal, preferably a *lats* knock-out mouse.

In a preferred embodiment, the invention provides a method for screening a potential therapeutic compound for activity in treating or preventing cancer. The potential therapeutic compound is administered to a recombinant non-human animal having at least one inactivated *lats* gene (i.e., a *lats* knock-out animal, preferably a *lats* knock-out mouse), preferably two inactivated *lats* genes (i.e., is homozygous for the inactivated *lats* allele). The size or progression of the cancer is then compared to that before the compound was added, or to a comparable recombinant animal without the administration of the compound, or to a normal, non-recombinant animal. A decrease in the size or progression of the cancer in the recombinant non-human animal after the administration of the compound as compared to the same animal prior to the administration or to another recombinant non-human animal not so administered or the standard size or progression of the cancer indicates that the compound has activity in treating or preventing cancer.

The screening method of the invention can be used to screen for potential therapeutic compounds for the treatment or prevention of any cancer, preferably a cancer or neoplastic disease that is caused by the *lats* knock-out mutation. As described in the Examples section, *infra*, *lats* knock-out mice are susceptible to ovarian stromal tumors and soft tissue sarcomas that metastasize to vital organs. Accordingly, in preferred embodiments, the invention provides methods for screening compounds useful in treating or preventing ovarian tumors and soft tissue sarcomas. *Lats* knock-out mutations in other animals or in other *lats* homologs may make the resulting knock-out animal susceptible to other types of neoplastic disease. The invention also contemplates use of these other *lats* knock-out animals to screen compounds for efficacy in treating or preventing the types of neoplastic diseases found in these *lats* knock-out animals. Additionally, compounds

effective to treat or prevent ovarian tumors and/or soft tissue sarcomas in *lats* knock-out animals may also be effective to treat or prevent other types of cancers and neoplastic disease. Thus, *lats* knock-out animals may be used to screen for compounds that have activity to treat or prevent these other types of cancers and neoplastic disease.

5 The invention also provides methods of screening compounds for efficacy in treating or preventing skin cancer. As demonstrated in Examples section, *infra*, exposure to carcinogens induced, at a high frequency, skin tumors in the *lats* knock-out mice. Many methods are known in the art for inducing skin carcinogenesis in animals (for review see DiGiovanni, 1992, Pharmac. Ther. 54:63-128). Generally, mouse skin tumors can be elicited by application of a carcinogenic dose of tumor initiator, *e.g.*, 600 to 800 nmole of a  
10 pure polycyclic aromatic hydrocarbon such as 9,10-dimethyl-1,2-benzanthracene (DMBA). Other tumor initiators include, but are not limited to, arylamines, carbamates, haloalkylethers, haloaromatics, lactones, nitro-aromatics, nitrosamides and ureas. Alternatively, and preferably, mouse skin tumors can be induced by an initial application of a single sub-carcinogenic dose of a tumor initiator, *e.g.*, DMBA, and then repeated doses or  
15 exposures to a tumor promoter, such as phorbol esters (*e.g.*, TPA), teleocidins, polyacetates, okadaic acid, calyculin A, palytoxin, and thapsigargin. Ultraviolet B (UVB) radiation, skin abrasion and skin wounding are also strong tumor promoters.

In a preferred embodiment, such skin tumors are induced by a two-step process comprising a single treatment with DMBA, preferably 50  $\mu$ l of a 0.5% DMBA solution in acetone, to the dorsal surface of the mouse 1 to 5 days after birth followed by repeated  
20 exposure to UVB irradiation, *e.g.*, exposures of approximately three times per week with an initial exposure of approximately 100 mJ/cm<sup>2</sup> per session, increasing the dosage by 10% per treatment (unless erythema or scaling occurs) to a maximum of 700 mJ/cm<sup>2</sup>, with an average of about 27 treatment sessions per mouse (Serrano et al., 1996, Cell 85:27-37).

Accordingly, in a preferred embodiment, the invention provides a method for  
25 screening a potential therapeutic compound for activity in treating or preventing skin cancer comprising administering the compound to a recombinant non-human animal in which one, preferably two, *lats* genes have been inactivated (*i.e.*, a *lats* knock-out animal) and in which recombinant non-human animal tumors have been induced by exposure to at least one carcinogen. The size or progression of the skin tumors are then compared before and after the administration of the compound. A reduction in the size or progression of the skin  
30 tumors in the recombinant non-human animal administered the compound as compared to same animal prior to administration of the compound or to the animal not so administered or to the standard size or progression of the skin tumors, indicates that the compound has activity in treating or preventing skin cancer. In another embodiment, the compound to be screened is administered by recombinantly expressing the compound in the recombinant  
35 non-human animal inactivated for the *lats* gene.

The administration of the compound to be tested can be carried out by any method known in the art, *e.g.*, orally, intravenously, intramuscularly, intraperitoneally,

subcutaneously, rectally, topically, etc. For the screening of compounds for efficacy in treating or preventing skin cancer, the compound is preferably applied topically.

After administration of the compound to be tested, the tumors, sarcomas, and other cancers can be evaluated by any diagnostic or histopathological method for detecting and evaluating tumors and cancers, for example, by visual inspection of the tumors (particularly for skin tumors), manual palpitation of tumors, biopsy or surgical removal of the tumor tissue and subsequent inspection, and sacrifice and dissection of the recombinant non-human animal. Morphological evaluation of tissue, either removed by biopsy or dissected from a sacrificed mouse, may be performed by fixing the tissue by any method known in the art, for example, in 10% neutral buffered formalin at 4°C, and subsequent dehydration, e.g., in ethanol. The fixed and dehydrated tissue may be embedded in paraffin and then sectioned, for example into 4-5 mm sections by any method known in the art. Sections can be stained, for example, with a standard stain, such as hematoxylin and eosin, for microscopic inspection.

Another aspect of the invention provides methods for screening potential therapeutic compounds for efficacy in treating or preventing diseases or disorders associated with pituitary dysfunction. *Lats* knock-out mice display a number of consequences of pituitary dysfunction, as described in the Examples section, *infra*. The methods of the invention can be used to screen compounds for efficacy in treating or preventing such pituitary dysfunctions as pituitary hyperplasia, fertility defects, such as defective ovulation, lack of breast development, abnormal reproductive cycles in females, LH hypogonadotropic hypogonadism, reduced levels of pituitary hormones, specifically LH, GH and PRL, and reduced growth and metabolic abnormalities caused by reduced GH levels. Therapeutics that are effective to treat one or more of these conditions associated with pituitary dysfunction may also be effective to treat or prevent other conditions, diseases or disorders associated with pituitary dysfunction.

In a preferred embodiment, potential therapeutic compounds to be screened for activity in treating or preventing diseases and disorders associated with pituitary dysfunction are administered to a recombinant non-human animal in which one or more chromosomal copies of the *lats* gene have been inactivated (*i.e.*, a *lats* knock-out animal, preferably and *lats* knock-out mouse). Levels of an indicator of pituitary function or dysfunction are then compared in the recombinant non-human animal before and after the compound was administered. In one embodiment, the compound to be screened is administered by recombinantly expressing the compound in the recombinant non-human animal having an inactivated *lats* gene.

Indicators of pituitary function that may be assayed include fertility, ovulation, the female reproductive cycle (*e.g.*, the estrus cycle), breast tissue development, growth or size of the animal, including weight, skeletal size, *e.g.*, of the skull and/or longitudinal bones, and organ weight, and serum levels of LH, GH and PRL. These indicators may be measured by any means known in the art for evaluating these indicators. For example,

fertility may be evaluated by attempting to mate an animal and determining whether conception occurred, measuring sperm count in male animals or detecting ovulation in female animals. The reproductive organ tissue may also be examined histopathologically (e.g., by fixing, sectioning and staining the tissue for inspection) for morphological defects, particularly in the testis, ovaries, and breast tissue. Whether the animal goes through an estrus cycle may be determined by observation of the animal. Hormone levels may be determined by any method known in the art, for example in serum samples by radio immunoassay using antibodies specific for the particular hormone. Lack of normal growth can be determined by measuring the animal e.g., the weight, size of the skull and/or longitudinal bones, or organ weight, during maturation.

### Candidate Therapeutics

Candidate therapeutics may come from any source of therapeutics known in the art. For example, these therapeutics can be proteins, nucleic acids (including anti-sense nucleic acids), antibodies, peptides, organic molecules, etc. In some instances, compounds may be screened first in *in vitro* assays to determine their potential as anti-cancer or anti-pituitary dysfunction therapeutics.

For example, chemical libraries may be screened for useful therapeutics. Exemplary libraries are commercially available from several sources (ArQule, Tripos/PanLabs, ChemDesign, Pharmacopoeia). Many diversity libraries suitable for use are known in the art and can be used to provide compounds to be tested according to the present invention.

Alternatively, libraries can be constructed using standard methods. Chemical (synthetic) libraries (Houghten et al., 1991, Nature 354:84-86; Lam et al., 1991, Nature 354:82-84; Medynski, 1994, Bio/Technology 12:709-710; Gallop et al., 1994, J. Medicinal Chemistry 37(9):1233-1251), recombinant expression libraries, or polysome-based libraries are exemplary types of libraries that can be used. Other examples include combinatorial libraries (Ohlmeyer et al., 1993, Proc. Natl. Acad. Sci. USA 90:10922-10926; Erb et al., 1994, Proc. Natl. Acad. Sci. USA 91:11422-11426; Houghten et al., 1992, Biotechniques 13:412; Jayawickreme et al., 1994, Proc. Natl. Acad. Sci. USA 91:1614-1618; Salmon et al., 1993, Proc. Natl. Acad. Sci. USA 90:11708-11712), organic diversity (e.g., nonpeptide) libraries (Bunin et al., 1994, Proc. Natl. Acad. Sci. USA 91:4708-4712) may be used.

Libraries of non-peptides, e.g., peptide derivatives (for example, that contain one or more non-naturally occurring amino acids) can also be used. One example of these are peptoid libraries (Simon et al., 1992, Proc. Natl. Acad. Sci. USA 89:9367-9371). Peptoids are polymers of non-natural amino acids that have naturally occurring side chains attached not to the alpha carbon but to the backbone amino nitrogen. Since peptoids are not easily degraded by human digestive enzymes, they are advantageously more easily adaptable to drug use.



### Therapeutic/Prophylactic Administration and Compositions

The invention provides methods of treatment (and prophylaxis) by administration to a subject of an effective amount of a Therapeutic of the invention. In a preferred aspect, the Therapeutic is substantially purified. The subject is preferably an animal, including but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, etc., and is preferably a mammal, and most preferably human. In a specific embodiment, a non-human mammal is the subject.

Various delivery systems are known and can be used to administer a Therapeutic of the invention, *e.g.*, encapsulation in liposomes (Langer, Science 249:1527-1533 (1990); Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353-365 (1989)), microparticles, microcapsules, recombinant cells capable of expressing the Therapeutic, receptor-mediated endocytosis (see, *e.g.*, Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), construction of a Therapeutic nucleic acid as part of a retroviral or other vector, etc. Methods of introduction include intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (*e.g.*, oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local.

In a specific embodiment where the Therapeutic is a nucleic acid encoding a protein Therapeutic, the nucleic acid can be administered *in vivo* to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, *e.g.*, by use of a retroviral vector (see U.S. Patent No. 4,980,286), or by direct injection, or by use of microparticle bombardment (*e.g.*, a gene gun; Biolistic, Dupont), etc. Alternatively, a nucleic acid Therapeutic can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a Therapeutic, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the Therapeutic is administered. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the Therapeutic, preferably in purified form, together with a suitable amount of carrier so as to

provide the form for proper administration to the patient. The formulation should suit the mode of administration.

In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection.

The amount of the Therapeutic of the invention which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges. However, suitable dosage ranges for intravenous administration are generally about 20-500 micrograms of active compound per kilogram body weight. Suitable dosage ranges for intranasal administration are generally about 0.01 pg/kg body weight to 1 mg/kg body weight. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems.

Suppositories generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations preferably contain 10% to 95% active ingredient.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention.

### Screening for Lats Agonists and Antagonists

Lats nucleic acids, proteins, and derivatives may be used in screening assays to detect molecules that specifically bind to lats nucleic acids, proteins, or derivatives and thus have potential use as agonists or antagonists of lats, in particular, molecules that thus affect cell proliferation and/or cdc2 activity, or molecules that promote or inhibit formation of lats-cdc2 complexes. In a preferred embodiment, such assays are performed to screen for molecules with potential utility as lead compounds for drug development, particularly as anti-cancer drugs. The invention thus provides assays to detect molecules that specifically bind to lats nucleic acids, proteins, or derivatives or bind to or interfere with the formation of lats-cdc2 complexes. For example, recombinant cells expressing *lats* nucleic acids can be used to recombinantly produce lats proteins in these assays, to screen for molecules that bind to a lats protein, and recombinant cells expressing *lats* and *cdc2* nucleic acids can be used to recombinant produce both lats and cdc2 proteins in these assays, to screen for molecules that bind to or inhibit formation of a lats-cdc2 complex. Molecules (*e.g.*, putative binding partners of lats) are contacted with the lats protein (or fragment thereof) under conditions conducive to binding, and then molecules that specifically bind to the lats protein or bind to or interfere with the formation of lats-cdc2 complexes are identified.

Similar methods can be used to screen for molecules that bind to lats derivatives or nucleic acids. Methods that can be used to carry out the foregoing are commonly known in the art.

By way of example, diversity libraries, such as random or combinatorial peptide or nonpeptide libraries can be screened for molecules that specifically bind to lats. Many libraries are known in the art that can be used, *e.g.*, chemically synthesized libraries, recombinant (*e.g.*, phage display libraries), and *in vitro* translation-based libraries.

Examples of chemically synthesized libraries are described in Fodor et al., 1991, *Science* 251:767-773; Houghten et al., 1991, *Nature* 354:84-86; Lam et al., 1991, *Nature* 354:82-84; Medynski, 1994, *Bio/Technology* 12:709-710; Gallop et al., 1994, *J. Medicinal Chemistry* 37(9):1233-1251; Ohlmeyer et al., 1993, *Proc. Natl. Acad. Sci. USA* 90:10922-10926; Erb et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:11422-11426; Houghten et al., 1992, *Biotechniques* 13:412; Jayawickreme et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:1614-1618; Salmon et al., 1993, *Proc. Natl. Acad. Sci. USA* 90:11708-11712; PCT Publication No. WO 93/20242; and Brenner and Lerner, 1992, *Proc. Natl. Acad. Sci. USA* 89:5381-5383.

Examples of phage display libraries are described in Scott and Smith, 1990, *Science* 249:386-390; Devlin et al., 1990, *Science*, 249:404-406; Christian, R.B., et al., 1992, *J. Mol. Biol.* 227:711-718; Lenstra, 1992, *J. Immunol. Meth.* 152:149-157; Kay et al., 1993, *Gene* 128:59-65; and PCT Publication No. WO 94/18318 dated August 18, 1994.

*In vitro* translation-based libraries include but are not limited to those described in PCT Publication No. WO 91/05058 dated April 18, 1991; and Mattheakis et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:9022-9026.

By way of examples of nonpeptide libraries, a benzodiazepine library (see *e.g.*, Bunin et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:4708-4712) can be adapted for use. Peptoid libraries (Simon et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:9367-9371) can also be used. Another example of a library that can be used, in which the amide functionalities in peptides have been permethylated to generate a chemically transformed combinatorial library, is described by Ostresh et al. (1994, *Proc. Natl. Acad. Sci. USA* 91:11138-11142).

Screening the libraries can be accomplished by any of a variety of commonly known methods. See, *e.g.*, the following references, which disclose screening of peptide libraries: Parmley and Smith, 1989, *Adv. Exp. Med. Biol.* 251:215-218; Scott and Smith, 1990, *Science* 249:386-390; Fowlkes et al., 1992, *BioTechniques* 13:422-427; Oldenburg et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:5393-5397; Yu et al., 1994, *Cell* 76:933-945; Staudt et al., 1988, *Science* 241:577-580; Bock et al., 1992, *Nature* 355:564-566; Tuerk et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:6988-6992; Ellington et al., 1992, *Nature* 355:850-852; U.S. Patent No. 5,096,815, U.S. Patent No. 5,223,409, and U.S. Patent No. 5,198,346, all to Ladner et al.; Rebar and Pabo, 1993, *Science* 263:671-673; and PCT Publication No. WO 94/18318.

In a specific embodiment, screening can be carried out by contacting the library members with a lats protein (or nucleic acid or derivative) immobilized on a solid phase and

harvesting those library members that bind to the protein (or nucleic acid or derivative). Examples of such screening methods, termed "panning" techniques are described by way of example in Parmley and Smith, 1988, Gene 73:305-318; Fowlkes et al., 1992, BioTechniques 13:422-427; PCT Publication No. WO 94/18318; and in references cited hereinabove.

5 In another embodiment, the two-hybrid system for selecting interacting proteins in yeast (Fields and Song, 1989, Nature 340:245-246; Chien et al., 1991, Proc. Natl. Acad. Sci. USA 88:9578-9582) can be used to identify molecules that specifically bind to a lats protein or derivative or that interfere with the formation of lats-cdc2 complexes. Additionally, the two hybrid system or co-immunoprecipitation of lats and cdc2 can be used as assays to  
10 screen for compounds that promote or inhibit formation of lats-cdc2 complexes.

In a preferred embodiment, the invention provides a method of screening for a molecule that modulates (*i.e.*, inhibits, antagonizes or promotes) directly or indirectly, the formation of a complex of lats and cdc2 proteins comprising measuring the levels of said complex formed from lats and cdc2 proteins in the presence of said molecule (optionally,  
15 purified) under conditions conducive to formation of the complex; and comparing the levels of said complex with the levels of said complex that are formed in the absence of said molecule, lower or higher level of said complex in the presence of said molecule indicates that the molecule modulates formation of said complex.

## 20 EXAMPLES

### EXAMPLE 1: Human Lats Modulates Cdc2/Cyclin A Activity

Using mammalian cell culture assays, we have found that lats is phosphorylated in a cell cycle-dependent manner and that it complexes with cdc2 in early mitosis. Lats  
25 associated cdc2 has no mitotic cyclin partner and no kinase activity for histone H1. Furthermore, we have found that *lats* mutant cells in *Drosophila* abnormally accumulate cyclin A. These biochemical observations indicate that lats is a negative regulator of cdc2/cyclin A, a finding supported by *in vivo* genetic data demonstrating that lats specifically interacts with cdc2 and cyclin A in *Drosophila*.

### 30 **Materials and Methods**

For yeast two-hybrid experiments, DNA encoding N-terminal h-lats (amino acid numbers 15 to 585 of the human lats protein sequence as depicted in Figure 12 (SEQ ID NO:2)), C-terminal h-lats (amino acid number 589 to 1130 of the human lats protein  
35 sequence as depicted in Figure 12 (SEQ ID NO:2)), and h-lats (amino acid numbers 15 to 1130 of the human lats protein sequence as depicted in Figure 12 (SEQ ID NO:2)) were cloned into pBTM116 (Bartel et al., Cellular Interactions in Development: A Practical Approach, ed. D. Hartley (Oxford University Press, Oxford, England (1993)). DNA

encoding human cdc2, CDK2, CDK4 and C-terminal h-lats were cloned in pACT (Durfee et al., 1993, Genes & Devel. 8:440-452). The constructs were transformed into yeast strain L40 (Vojtek et al., 1993, Cell 74:205-214). The transformants were tested for growth on SD his-ura-trp-leu medium (Bio101, California) and for  $\beta$ -galactosidase activities (Vojtek et al., 1993, Cell 74:205-214).

For baculovirus experiments, full-length *h-lats* cDNA was cloned into the vector pBacPAK8 and baculovirus were produced according to the protocols provided by Clontech. IPLB-Sf21 cells were co-infected with equal amounts of h-lats and human cdc2-baculoviruses and were harvested 62 hours after infection for immunoprecipitation and immunoblot assay.

### Fly Genetics

The full-length *h-lats* cDNA was cloned into the vector pCaSpeR-hs (Tummel and Pirrott, 1992, *Drosophila* Information Service 71:150). Multiple transformant lines were obtained and used in rescue experiments with *lats*<sup>es32</sup>, *lats*<sup>e26-1</sup>, *lats*<sup>a1</sup>, and *lats*<sup>X1</sup> alleles. Expression of *hs-h-lats* was induced as described in Xu et al. (1995, Development 121:1053-1063) -- incubation at 37°C for one hour every day until eclosure. Since the induction of *h-lats* requires heat-shock treatments, X-ray irradiation was used to induce mitotic clones in *y w P[hs-h-lats]/y w; P[FRT]82B lats<sup>X1</sup>/P[y+]96E* animals. The rest of the *lats* mutant mitotic clones were induced and labeled according to Xu and Rubin (1993, Development 117:1223-1237). The *Drosophila lats* and *h-lats* cDNAs were cloned into the pGMR vector to generate multiple transformant lines (Hay et al., 1994, Development 120:2121-2129). Besides the mutations mentioned above, *cdc2c*<sup>E136E</sup> (a gift of Helena Richardson), *cycA*<sup>neo114</sup>, *Df(2R)59A-B* were used for *cdc2c*, *cyclin A* and *cyclin B*, respectively.

### Tissue Culture

HeLa cells were synchronized at different cell cycle stages by various treatments as described by Knehr et al. (1995, Exp. Cell. Res. 217:546-553). Briefly, cells were arrested at G1 by thymidine and hydroxyurea treatment; at S phase by thymidine double block (incubation in the presence of 2 mM thymidine for 24 hours, followed by recovery in the absence of thymidine for 12 hours, followed by another incubation in 2 mM thymidine for 14 hours) plus a 4 hour incubation in medium without thymidine; and at G2 by thymidine double block plus an 8 hour incubation in medium without thymidine. To arrest cells in M phase, cells were treated with 0.1  $\mu$ g/ml nocodazole (Sigma) for 12 hours and mitotic cells were shaken off of the flask and washed twice with cold DMEM without serum. These mitotic cells were then resuspended in fresh warm medium without nocodazole and incubated in suspension at 37°C. Cells were harvested at various time points after removal of nocodazole (herein "ARN") for further analysis. CHO cells were grown in  $\alpha$ -MEM medium plus 7% FBS and IPLB-Sf21 cells were grown in sf-900 II SFM plus 10% FBS.

### Antibodies and Immunochemistry

Anti-human *lats* rat monoclonal and rabbit polyclonal antibodies were raised against a GST-N-h-Lats (GST fused to the N-terminal portion of *lats*, *i.e.*, consisting of amino acids 15-585 of the human *lats* amino acid sequence as depicted in Figure 12 (SEQ ID NO:2) fusion protein. Anti-human *cdc2* (#sc 054), anti-human cyclin B (#sc 245), anti-human cyclin A (#sc 239) monoclonal antibodies were purchased from Santa Cruz Inc. Rabbit polyclonal anti-*Drosophila* Cyclin A and B antibodies were gifts of David Glover. Monoclonal mouse anti-BrdU antibodies (#347580) were purchased from Becton Dickinson and monoclonal mouse anti-c-myc antibodies (#OP 10) were purchased from Oncogene Sciences. Propidium iodide (Sigma) was used as a DNA marker.

HeLa, CHO, or IPLB Sf21 cells were lysed in TG buffer (1% Triton, 10% glycerol) (Sun et al., 1996, Genes & Devel. 10:395-406) plus freshly added proteinase inhibitors (1 mM PMSF, 10 µg/ml leupeptin, 1 µg/ml pepstatin, and 2 µg/ml aprotinin). The lysates were centrifuged at 50 K, 4°C for 12 minutes. Supernatants were pre-cleaned by incubating with protein G-agarose. Immunoprecipitation and western blots were performed by the procedures described by Sun et al. (1996, Genes & Devel. 10:395-406). Western blots were visualized by enhanced chemiluminescence (Amersham). Whenever necessary the blots were stripped following the procedure described by Edgar et al. (1994, Genes & Devel. 8:440-452). Calf Intestinal Phosphatase (CIP) treatments were carried out as described in Sun et al. (1996, Genes & Devel. 10:395-406).

### H1 Kinase Assay

HeLa cell lysates (50 minutes ARN) were precleaned by incubation in protein G-agarose. Immunoprecipitates were washed three times with TG buffer and twice with 1X kinase buffer (50 mM Tris-HCl 7.5, 10 mM MgCl<sub>2</sub>, 5 mM EGTA, 2 mM DTT) without DDT. The kinase assay was carried out on ice for 10 minutes in 35 µl of 1X kinase buffer containing 15 µCi of γ-<sup>32</sup>P ATP, 1.6 µg of histone H1, and 1.5 µM ATP. The kinase activities were measured by quantifying the intensities of histone-H1 phosphorylation using a PhosphorImager (Molecular Dynamics). The amounts of *cdc2* in the immunoprecipitates were determined by anti-*cdc2* immunoblotting and densitometer scanning (Molecular Dynamics). The kinase assay experiments were repeated three times.

## Results

### Human Lats Can Functionally Replace its Fly Counterpart

Sequence conservation suggested that human *lats* could be a functional homolog of *Drosophila* *lats*. To test this, the human *lats* cDNA was introduced into the *Drosophila* genome under the control of the heat shock-inducible promoter (*hs-h-lats*) (Lis et al., 1983, Cell 35:403-410) and expressed the transgene was expressed under the conditions previously established for rescue using the fly *lats* gene (Xu et al., 1995, Development 121:1053-1064; and PCT Publication WO 96/30402, published October 3, 1996). In

mosaic flies, clones of cells mutant for *lats* undergo extensive overproliferation and develop into large tumors in various tissues (Figure 1A and Xu et al., 1995, Development 121:1053-1064). Expression of human *lats* completely suppressed tumor formation in *lats* mosaic flies (Figures 1B-D). Instead, the *lats* mutant cells (genetically marked as *yellow-* cells) in these human *lats*-expressing mosaic animals developed into normal adult structures (Figures 1C and D). The ability of the human gene to support normal fly development was further examined. Expression of the human *lats* transgene in homozygous *lats* mutant *Drosophila* rescued all developmental defects including embryonic lethality found in homozygous *lats* mutants. Furthermore, the extent of phenotypic rescue correlated with the level of human *lats* expression. Complete phenotypic rescue required daily induction of human *lats*, and leaky expression of *lats* controlled by the heat shock promoter at 25°C resulted in partial suppression of the *lats* mutant overproliferation phenotype (Figures 1E and F). These data demonstrate that human *lats* is an authentic homolog of the *Drosophila* *lats* tumor suppressor.

#### **Lats is phosphorylated in a cell cycle-dependent manner**

To further explore the function of *lats*, the biochemical properties of the human *lats* protein were examined. *Lats* immunoprecipitated from HeLa cells had two major migrating forms (Figure 2A, lane 6). The slow-migrating form of *lats* was converted into the fast-migrating form after the proteins were incubated with calf intestine alkaline phosphatase (CIP) (Figure 2A). Addition of a phosphatase specific inhibitor,  $\beta$ -glycerophosphate, to the phosphatase reaction blocked this conversion (Figure 2A, lanes 5 and 10). These results indicate that the slow-migrating form is phosphorylated *lats*, while the fast-migrating form is dephosphorylated *lats*.

*Lats* immunoprecipitated from cells at different mitotic stages displayed varying amounts of the two forms (compare lanes 1 and 6 of Figure 2A), suggesting that the phosphorylation state of *lats* may oscillate with the cell-cycle. To verify this possibility, *lats* proteins were immunoprecipitated from extracts of HeLa cells at G0, G1, S, and G2 phases, and different time points during mitosis (minutes after removal of nocodazole (ARN) block) (Knehr, et al., 1995, Exp. Cell Res. 217:546-553). DAPI staining was used to verify the cell cycle progression. All *lats* protein was phosphorylated at late prophase (0 minutes ARN; Figure 2B), and remained phosphorylated through metaphase (50 minutes ARN; Figures 2B and C). Dephosphorylated *lats* could be detected when cells in the culture begin to enter anaphase (75 minutes ARN; Figures 2B and C), and by the start of telophase (100 minutes ARN) most of the *lats* protein was dephosphorylated (Figures 2B and C). In late mitosis, G1, S, G2 or G0 phase, *lats* molecules were in the dephosphorylated form (Figure 2B). These observations strongly suggest that the *lats* protein undergoes two major phosphorylation changes during the cell cycle. At the G2/M boundary or in early prophase, *lats* is phosphorylated, and *lats* becomes dephosphorylated at the metaphase/anaphase boundary or in early anaphase.

### Lats complexes with cdc2 during mitosis

*Lats* mutant cells in *Drosophila* mosaic for the *lats* mutation do differentiate, indicating that mutations in *lats* do not block cellular differentiation in general (Figures 1G and H). The *lats* mutant overproliferation phenotype and cell cycle-dependent phosphorylation of lats suggest that the protein could be directly involved in the regulation of the cell cycle.

Immunoprecipitation experiments were carried out to examine whether lats proteins complex with known cell cycle regulators. Interestingly, cdc2 was found to co-immunoprecipitate with lats in mitotic cells (Figures 3A and B). The co-precipitation of lats and cdc2 was confirmed in both murine and human cells using several polyclonal and monoclonal anti-human lats antibodies (as described above in the Materials and Methods section, *supra*). Although similar amounts of lats were immunoprecipitated from HeLa cells during different stages of mitosis (Figure 3B, upper panel), the amount of co-precipitated cdc2 varied (Figure 3B, lower panel). Co-precipitated cdc2 was most abundant at early mitosis (0 and 50 minutes ARN; Figure 3B; also see Figures 2B and C for cell cycle progression). The amount of co-precipitated cdc2 then progressively decreased as the cell cycle progressed (Figure 3B--lanes 100', 150', and 200'). No cdc2 co-immunoprecipitation could be detected in quiescent cells in G0 (Figure 3A and Figure 2B). The difference in the amount of co-precipitated cdc2 cannot be attributed to changes of cdc2 levels during the cell cycle, since it has previously been shown that the cdc2 protein is maintained at a nearly constant level in cycling cells (Dalton, 1992, EMBO J. 11:1797-1804; McGowan et al., 1990, Mol. Cell. Biol. 10:3847-3851). Indeed, equal amounts of cdc2 protein were precipitated from the 50 minutes and 200 minutes ARN extracts when anti-cdc2 antibodies were used.

The interaction of lats and cdc2 proteins was studied by expressing human lats and cdc2 proteins in the baculovirus expression system. The baculovirus-expressed cdc2 and lats proteins could be co-immunoprecipitated using either anti-human lats or anti-cdc2 antibodies (Figure 3C). This result suggests that the *in vivo* lats/cdc2 complex may result from direct binding of the two proteins.

The interaction between lats and cdc2 was also examined using the yeast two-hybrid assay (Fields and Song, 1989, Nature 340:245-245; Finley and Brent, in DNA Cloning 2, Rickwood and Hames, eds (Oxford University Press (Oxford, 1995))). Consistent with the co-immunoprecipitation results, full length lats and the N-terminal region of lats interacted with cdc2 in the assay (Figure 3E). Since the C-terminal kinase domain of lats did not interact with cdc2, these results indicate that lats associates with cdc2 through its N-terminal domain. Furthermore, neither full length lats nor the N-terminal region of lats showed any interaction with two G1 cell cycle kinases, CDK2 and CDK4 (Figure 3E), indicating that the association between lats and cdc2 is specific.



### **The lats/cdc2 complex is inactive for H1 kinase activity**

We examined whether the lats/cdc2 complex has any kinase activity on the substrate histone H1. Lats/cdc2 and cdc2/cyclin B complexes were immunoprecipitated separately from 50 minute ARN HeLa cell extracts using either anti-human lats or anti-cyclin B monoclonal antibodies and assayed for histone H1 kinase activities. In contrast to the cdc2/cyclin B complex, the lats/cdc2 complex showed no detectable kinase activity for histone H1 (Figure 3D). Densitometer readings indicated that the H1 kinase activity of the lats/cdc2 complex does not differ from the background control and is at least 25 fold lower than the kinase activity of the cdc2/cyclin B complex. These results indicate that cdc2 molecules associated with lats are inactive or have dramatically reduced mitotic kinase activity.

The lack of H1 kinase activity in the lats-associated cdc2 could be due to the inhibition of the kinase activity of the cdc2/cyclin complex by lats. Alternatively, the lats/cdc2 complex may lack cyclin A and B which are the indispensable subunits for cdc2 kinase activity (Draetta et al., 1989, Cell 56:829-838; Solomon et al., 1990, Cell 63:1013-1024). Neither cyclin A nor cyclin B proteins could be detected in the lats/cdc2 immunocomplex when probed with anti-cyclin A and B antibodies (Figure 3B), indicating that lats modulates cdc2 activity in a way different from that of the known cyclin dependent kinase inhibitors (CDIs) (Sherr, 1996, Science 274:1672-1677; Harper, 1997, Cancer Surveys 29:91-107).

### **Lats genetically interacts with cdc2 and cyclin A during *Drosophila* development**

In *Drosophila*, cdc2 also complexes with cyclin A or B (Knoblich et al., 1994, Cell 77:107-120). We examined the potential genetic interactions between lats, cdc2, cyclin A and cyclin B in *Drosophila*. Animals heterozygous for the strong *cdc2* allele, *cdc2<sup>B47</sup>*, or homozygous for the temperature sensitive *cdc2<sup>ts</sup>* mutation at permissive temperature are viable and morphologically normal (Clegg et al., 1993, Genome 36:676-685). The *lats<sup>P8</sup>* mutation causes late pupal lethality in homozygous mutants (Figure 4A), and reducing cdc2 activity in *lats<sup>P8</sup>* homozygotes by introducing one copy of a *cdc2* mutant allele (*cdc2<sup>B47</sup>* or *cdc2<sup>ts</sup>/+*; *lats<sup>P8</sup>/lats<sup>P8</sup>*) was sufficient to rescue the lats-associated lethality (Figure 4B). Furthermore, the overproliferation phenotype of *lats<sup>P8</sup>* adult appendages were also suppressed. Rescued animals had near-wild type eyes in comparison to the overproliferated, large, rough eyes of the *lats<sup>P8</sup>* mutants (Figures 4C and D). Reducing cdc2 activity also suppressed the giant larvae/pupae and disc overproliferation phenotypes of the *lats<sup>E26-1</sup>* animals (Figures 4E and F). The *Drosophila* CDK2 homolog, cdc2c complexes with Cyclin E (Lehner and O'Farrell, 1990, Cell 61:535-547; Knoblich et al., 1994, Cell 77:107-120). Consistent with the result of the yeast two-hybrid assay for human lats and CDK2 (Figure 3C), inactivation of one copy of the *cdc2* gene did not modify the phenotypes of the *lats<sup>P8</sup>* mutant animals. We further examined the potential genetic

interaction between *lats*, *cyclin A* and *cyclin B* in *Drosophila* as described above. Interestingly, while cyclin B did not interact with *lats*, cyclin A behaved similarly to *cdc2*. Inactivation of one copy of the cyclin A gene resulted in almost identical phenotypic suppression of the *lats*<sup>P8</sup> and *lats*<sup>E26-1</sup> mutants as did the *cdc2* mutants (data not shown). Thus, the specific genetic interactions between *lats*, *cdc2*, and cyclin A confirms the biochemical data indicating that *lats* regulates cell proliferation by negatively modulating *cdc2*/cyclin A activity.

While *cdc2* protein remains at a constant level during the cell cycle (Dalton, 1992, EMBO J. 11:1797-1804; McGowan et al., 1990, Mol. Cell. Biol. 10:3847-3851), cyclin A and B are degraded when the *cdc2*/cyclin complexes are inactivated (Draetta et al., 1989, Cell 56:829-828; Murray et al., 1989, Nature 339:280-286; King et al., 1994, Cell 79:563-571).

Accordingly, whether *lats* inactivation leads to accumulation of cyclin A and B was also determined. In the *Drosophila* eye imaginal disc, cyclin A and B are detected in cells anterior to the morphogenetic furrow (MF) as well as in a stripe of cells posterior to the MF which are undergoing the last round of cell division (the second mitotic wave) (Thomas et al., 1994, Cell 77:1003-1014). In clones of *lats* cells located anterior to the MF, anti-cyclin A or B antibody staining did not detect any obvious changes in levels of the two proteins (Figures 5A-N; Whitfield et al., 1990, EMBO J. 9:2563-2572), which might be due to the fact that cells in this region already accumulate high levels of both cyclin A and B. Cells in the MF are synchronized in G1 and so are the cells posterior to the MF which are differentiating into neurons. Both populations of the G1 cells have no detectable cyclin A or B (Thomas et al., 1994, Cell 77:1003-1014; Figures 5B and H). However, in clones of *lats*<sup>+</sup> cells in the MF and in the region posterior to the MF, a high level of cyclin A was detected (Figures 5A-E). Interestingly, in *lats*<sup>-/-</sup> cells, cyclin A was still degraded in cells at late mitosis (Figure 5F), indicating that *lats* affects limited aspects of the cell cycle. Finally, consistent with the genetic interaction results, *lats*<sup>-/-</sup> clones did not cause obvious changes in the staining pattern of cyclin B (Figures 5G and H). These results provide direct evidence indicating that inactivation of *lats* causes overproliferation by deregulating *cdc2*/cyclin A activity.

The role of *lats* in cell cycle regulation was examined by overexpressing *lats* in the developing eye imaginal disc. When *Drosophila* and human *lats* cDNAs were expressed in cells in and posterior to the MF under the direction of the GMR promoter (Hay et al., 1994, Development 120:2121-2129), they exhibited similar phenotypes. Adult eyes from these animals were smaller than wild type and had irregular architecture with missing bristles, a phenotype reminiscent to that of overexpressing p21 under the same promoter (*GMR-p21*; de Nooij and Hariharan, 1995, Science 270:983-985) (Figures 5I and J). Sections of GMR-*lats* adult retinas also revealed a phenotype identical to that observed in retinas from flies transformed with *GMR-p21*. While almost all of the ommatidia contained the full complement of photoreceptor cells, many pigment cells were missing (Figure 5K). This

phenotype suggested that, like overexpression of p21, overexpression of lats blocked the last round of cell division in the developing eye disc. Further examination of the GMR-lats eye discs revealed that in the region of second mitotic wave there was an accumulation of cells with intense propidium-iodide staining, indicating that the cells were tetraploid, which was followed immediately posteriorly by apoptotic cells with fragmented nuclei (Figure 5L). In contrast to the overexpression of p21 which blocked entry into S phase, overexpressing lats arrested cells at G2/M or M phase. Consistent with this conclusion, BrdU labeling experiments revealed that the S phase of the cells in the second mitotic wave did occur in GMR-lats eye discs (Figures 5M and N).

## 10 Discussion

### The lats molecules are a novel family of conserved proteins

Expression of human lats under the same conditions used for rescue by the fly gene completely suppressed tumor formation in *lats* mosaic flies and rescued all developmental defects in *lats* homozygous mutant *Drosophila*. These experiments provide definitive evidence for functional conservation among the *lats* genes, indicating that human lats can perform all functions that are normally provided by the fly protein.

### A model for lats function

Our biochemical and genetic data support the hypothesis that lats is a negative regulator of cdc2/cyclin A. Cdc2 co-immunoprecipitated with lats using either HeLa and CHO cell extracts or baculovirus-expressed proteins. Cdc2 also interacted with lats in yeast two-hybrid assays. Moreover, lats-associated cdc2 has no cyclin A or B subunit and no histone H1 kinase activity. In *Drosophila*, *lats* mutant cells abnormally accumulated cyclin A. Genetic data in *Drosophila* demonstrate that the overproliferation and lethality phenotypes of *lats* mutants can be suppressed by mutations in *cdc2* and *cyclin A* genes. The genetic interaction between lats, cdc2, and cyclin A is highly specific. While animals heterozygous for cdc2 or cyclin A did not display any defects, removal of one copy of either gene was sufficient to dominantly suppress the *lats* mutant phenotypes. Such a genetic interaction was not observed in hundreds of genes examined, including other positive cell cycle regulators such as *cdc2c*, *cyclin B*, and *dE2F*.

The lats kinase domain contains all 11 subdomains previously found in other protein kinases (Hanks et al., 1988, Science 241:42-45), suggesting that it is an active protein kinase. However, lats alone and lats/cdc2 complex do not appear to have any autophosphorylation activity or phosphorylation activity for cdc2 and histone H1. Yeast two-hybrid experiments showed that the N-terminal region of lats interacted with cdc2 much more strongly than did full-length lats (Figure 3E). This result indicates that the C-terminal kinase domain of lats has a negative effect on the binding between the lats N-terminal region and cdc2.

The association of *lats* with *cdc2* is directly correlated with its state of phosphorylation (compare Figures 2B and 3B). During early mitosis, *lats* is phosphorylated and associates with *cdc2*. At G0, *lats* is dephosphorylated and fails to associate with *cdc2*. Furthermore, the transition of the *lats* phosphorylation state during mitosis correlates with a change in its ability to bind to *cdc2*. Phosphorylation is a common mechanism that regulates protein activities during the cell cycle (Hunter, 1995, Cell 80:225-236).

### Cell-cell communication mechanism regulating cell proliferation

Many tumor suppressors probably evolved to play important regulatory roles during development. The study of the normal developmental functions of a tumor suppressor is essential to our understanding of the mechanisms of tumorigenesis. A growing body of evidence suggests that proliferating cells in a developing *Drosophila* imaginal disc communicate to maintain a constant disc size, and that *lats* plays an important role in this process. Using mutations such as *Minute* and *dE2F* in *Drosophila*, it has been shown that, from a young mosaic disc containing cells of different genotypes, the number of progeny cells from a given parental cell can vary dramatically in a mature disc, while the overall size of the mature disc is unaffected (Simpson, 1979, Devel. Biol. 69:182-193; Simpson and Morata, 1981, Devel. Biol. 85:299-308; Brook et al., 1996, EMBO J. 15:3676-3683). Imaginal discs can also undergo regeneration when a small region of a disc is surgically removed, a phenomenon similar to liver regeneration in mammals (French et al., 1976, Science 193:969-981; Meinhardt, 1994, Bioessays 16:627-632; Michalopoulos and DeFrances, 1997, Science 276:60-66). Consistent with the notion that proliferation is regulated by local cell interaction, it has been shown that DNA replication and mitosis in growing discs occur in small, non-clonal clusters of cells throughout the disc (Adler and MacQueen, 1981, Exp. Cell Res. 133:452-456; Milan et al., 1996, Proc. Natl. Acad. Sci. USA 93:11687-11692; Milan et al., 1996, Proc. Natl. Acad. Sci. USA 93:640-645). Furthermore, young discs transplanted into adult hosts grow until the disc reaches its normal size, indicating that such a size control mechanism is an intrinsic property of the cells in each disc (Bryant in The Genetics and Biology of *Drosophila*, Vol. 2c, Ashburner and Novitski, eds. (Academic Press, New York, 1978). Transplantation experiments in mice revealed a similar size control phenomenon with anlagen of some vertebrate organs (Leitina et al., 1971, Transplantation 11:499-502).

*Lats* mutations dramatically disrupt the size and shape of discs in *Drosophila*. Clones of *lats* mutant cells in mosaic discs overproliferate to form massive outgrowths that are sometimes larger than the mature discs themselves, and animals homozygous for many *lats* alleles also have dramatically overgrown discs (Figures 1A and F; Xu et al., 1995, Development 121:1053-1063). These *lats* phenotypes indicate that an inhibitory cell-cell communication mechanism has been disrupted and suggest that the *lats* protein could be a component of this mechanism regulating cell proliferation. The overproliferation phenotype of *lats* behaves in a cell autonomous fashion: inactivating *lats* causes mutant cells

to overproliferate (Xu et al., 1995, Development 121:1053-1063). Furthermore, in mosaic discs containing *lats* mutant clones, there is an overproliferation of *lats* mutant cells as well as a reduction in the number of wild type cell. These observations are consistent with a regulatory mechanism where *lats* mutant cells are able to send signals inhibiting cell proliferation but are defective in receiving such signals.

While the mammalian cdc2/cyclin A complex is involved in G2/M regulation (Hamaguchi et al., 1992, J. Cell Biol. 117:1041-1053; Hunter and Pines, 1994, Cell 79:573-582), *Drosophila* cdc2/cyclin A functions at the G1/S phase transition in addition to the G2/M phase transition. Ectopic activation of cdc2/cyclin A by overexpressing cyclin A in G1 arrested cells can drive the G1/S transition and induce S phase in cells lacking cyclin E (Dong et al., 1997, Genes & Devel. 11:94-105; Sprenger et al., 1997, Curr. Biol. 7:488-499). This G1/S activity is greatly enhanced when both cyclin A and an activated form of cdc2 are overexpressed. In *roughex* (*rux*) mutants, cells accumulate cyclin A in early G1 and progress into S phase precociously (Thomas et al., 1994, Cell 77:1003-1014; Thomas et al., 1997, Genes & Devel. 11:1289-1298). Loss of fizzy-related (*fzr*), a cdc2-related fly gene, results in accumulation of mitotic cyclins in G1 cells and causes progression through an extra division cycle in the embryonic epidermis (Sigrist and Lehner, 1997, Cell 90:671-681). These observations have shown that in *Drosophila* extra cdc2/Cyclin A activity can cause overproliferation. Consistent with these observations, we find that Cyclin A is abnormally accumulated in *lats* mutant cells (Figures 5A-N) and *lats* phenotypes can be suppressed by *cdc2* and *cyclin A* mutations. Several aspects of the *lats* phenotype are unique. First, *lats* mutants deregulate cdc2/cyclin A activities which affects both the G1/S and G2/M transitions. Second, while mutants such as *rux*, *fizzy*(*fzy*), and *fzr* accumulate multiple mitotic cyclins and thus affect activities of several cdc2/cyclin complexes (Thomas et al., 1994, Cell 77:1003-1014; Dawson et al., 1995, J. Cell Biol. 129:725-737; Sigrist and Lehner, 1997, Cell 90:671-681), *lats* mutants appear to only affect cdc2/cyclin A. Finally, in *lats* mutant cells, cyclin A is degraded at late mitosis (Figure 5F), further indicating that many aspects of the cell cycle are normal in *lats* mutants. These properties distinguish *lats* mutants from genetic alterations that affect multiple CDK/cyclin complexes or that abnormally activate CDK/cyclin at a single cell-cycle stage or throughout the entire cell cycle, and provide an explanation for the extensive overproliferation phenotype of the *Drosophila lats* mutants.

The data provided herein indicate that cdc2/cyclin A activity is negatively regulated by the *lats* protein. Yeast two-hybrid assays show that *lats* specifically interacts with cdc2 but not other CDKs. Genetic data in *Drosophila* also show that *lats* interacts with cdc2 but not the fly CDK2 homolog, cdc2c. Given that p16- and p21-like CDK inhibitors have not been found for cdc2, it is possible that cdc2 and the rest of the CDKs are negatively regulated by different families of proteins. Alternatively, the activity of each CDK could be modulated by both types of negative regulators. In both flies and mammals, cdc2/cyclin A is inactivated during early mitosis by degradation of cyclin A, while degradation of cyclin B

occurs later at the metaphase/anaphase transition (Minshull et al., 1990, EMBO J. 9:2865-2875; Whitfield et al., 1990, EMBO J. 9:2563-2572). The mechanism of such differential inactivation of cdc2/cyclin is unknown. Our data indicate that lats specifically modulates cdc2/cyclin A activity but not cdc2/cyclin B activity: cyclin A but not cyclin B mutants interact with lats genetically; *lats* mutant cells abnormally accumulate cyclin A but not cyclin B (Figures 5A-N).

Finally, overexpression of cdc2 and cyclin A has been reported in multiple types of human tumors (Wang et al., 1990, Nature 343:555-557; Keyomarsi and Pardee, 1993, Proc. Natl. Acad. Sci. USA 90:1112-1116; Arany et al., 1994, Surg. Onc. 3:153-159). Negative regulators of CDK/Cyclins (*e.g.*, p16) have been shown to function as tumor suppressors in mammals (Serrano et al., 1996, Cell 85:27-37). The biochemical and genetic data for lats provided herein suggest that lats would behave as a tumor suppressor in mammals.

## **EXAMPLE 2: Mice Deficient for Lats Develop Soft Tissue Sarcomas, Ovarian Tumors and Pituitary Dysfunction**

### **Materials and Methods**

#### **Generation of *lats*<sup>-/-</sup> mice**

Mouse *lats* genomic DNA was isolated by screening a 129 library (Stratagene) using a mouse *lats* cDNA as a probe. A SalI fragment from the cDNA was subcloned into a pBS vector. We cleaved this construct at the EcoRV site (Figure 6B), and inserted a 1.8 kb fragment encoding PGK-neo. We subsequently digested with BamHI and XhoI and inserted a 3 kb PGK-TK gene cassette.

D3 embryonic stem (ES) cells were electroporated with the SfiI linearized vector, and selected in 0.3 mg/ml G418 and 2  $\mu$ M ganciclovir media for incorporation of the vector. The ES cell clones analyzed underwent homologous recombination at the lats locus (Figure 6B). For genotyping, genomic DNA from the ES cells was digested with BamHI and EcoRV and analyzed by Southern blotting using the BamHI-EcoRI probe from the vector (Figure 6C). The double digest of the wild type allele generates a 3.5 kb fragment that hybridizes to the probe, while double digest of the disrupted allele generates a 5.8 kb fragment that hybridizes to the probe. *Lats* heterozygous ES cells were microinjected into CS7BL/6 blastocysts which were transplanted into uteri of pseudopregnant ICR mice. Chimeric male progeny were crossed to CS7BL/6 females. Germline transmission of the disrupted allele was detected in agouti progeny by Southern blotting.

#### **Cell culture and protein analysis**

Proteins were extracted from whole-cell lysates of *lats*<sup>-/-</sup> mouse embryonic fibroblasts (MEFs) derived from 13-days post-coitum mouse embryos, separated using SDS-PAGE, transferred and probed with rabbit polyclonal anti-lats antibody, followed by enhanced chemiluminescence detection (Amersham).

### Histopathological examinations

For morphological evaluation, tissues were fixed in 10% neutral buffered formalin at 4°C overnight, dehydrated with ethanol, embedded in paraffin, and sectioned into 4 to 5 mm sections. Paraffin sections were prepared by standard procedures and stained with hematoxylin and eosin.

### Gonadotropin treatment

Mice were injected intraperitoneally with FSH administered in the form of 5 IU of pregnant mare serum gonadotropin (Sigma). 44-46 hours later, mice were injected intraperitoneally with LH in the form of 5 IU of human chorionic gonadotropin (Sigma).

### Pituitary hormone measurements

We used 20 *lats*<sup>-/-</sup> and 20 *lats*<sup>+/-</sup> age, sex, and estrus cycle matched females and males for these analyses. Mouse serum levels of PRL, LH, GH, FSH, and TSH were determined in pooled serum samples by double antibody radioimmunoassays (RIAs). These sensitive, specific mouse pituitary hormone RIAs were developed by A. F. Parlow, and are distributed to the scientific research community via the National Hormone & Pituitary Program of NIDDK, NIH (see <http://www.humc.edu/hormones>).

### UVB and DMBA tumorigenic treatments

UVB and DMBA treatments were performed as described by Serrano et al. (1996, Cell 85:27-37). Briefly, skin tumors were induced by first applying a single dose of 9,10-dimethyl-1,2-benzathralene (DMBA; 50 µl of an 0.5% solution in acetone) to the dorsal surface of the mouse 1 to 5 days after birth. This treatment was followed by exposure to ultraviolet B (UVB) irradiation approximately three times per week for, on average, 27 treatments, with an initial exposure of approximately 100 mJ/cm<sup>2</sup>, increasing the dosage by 10% per treatment to a maximum of 700 mJ/cm<sup>2</sup>.

## Results and Discussion

### Targeted Disruption of the mouse *lats* gene

A 17.5 kilobase *lats* genomic clone obtained from a mouse 1295V library was used to construct a targeting vector for homologous recombination by positive-negative selection (Mansour et al., 1988, Nature 336:348-352; Capecchi, 1989, Science 244:1288-1292) as shown in Figure 6B. A PGK-neo cassette was inserted in inverse orientation into an exon of the *lats* clone resulting in the removal of amino acid sequence corresponding to amino acids 756-1130 of human *lats* (Figure 6A). We electroporated D3 embryonic stem cells (Gossler et al., 1986, Proc. Natl. Acad. Sci USA 83:9065-9069) with the *lats*-neo construct; single clones resistant to G418 and ganciclovir were screened by Southern blot hybridization using a mouse *lats* probe 5' to the portion of the *lats* gene contained in the targeting vector (Figure 6B). Restriction enzyme digestion of the wild-type *lats* locus with

BamHI and EcoRV generated a 3.5 kb fragment, while the correctly targeted, disrupted locus generated a 5.8 kb fragment (Figures 6B and C). A targeting frequency of approximately 1 in 100 was observed. Male chimeras transmitted the targeted *lats* allele through the germline, as demonstrated by Southern blot analysis of tail DNA.

Immunoprecipitation and western blotting of *lats*<sup>-/-</sup> mouse embryonic fibroblast lysates (derived from 13.5 days post-coitum embryos (dpc)) with polyclonal anti-human *lats* antibody confirmed the absence of *lats* protein in homozygous null embryos (Figure 6D).

### Growth and viability of *lats*<sup>-/-</sup> mice

When assessed at 3 weeks of age, the number of *lats* mutant animals was drastically lower than expected; 8% of pups genotyped from double heterozygote matings were *lats*<sup>-/-</sup>, significantly lower than the expected frequency of 25%. Just before birth at embryonic day 18.5 (dpc), however, live homozygous embryos were found at the frequency predicted by Mendelian law (25%). The majority of homozygotes died within the first day of life, and their death was associated with internal hemorrhage into vital organs. The reason that some *lats*<sup>-/-</sup> mice survived while most did not is uncertain, but the mixed genetic background (strains 129 and C57BL/6) of the mice could be contributory. The weight of homozygous embryos at birth was approximately 70% of that of wild-type embryos (*lats*<sup>+/+</sup>, 1.5 ± 0.3 g; *lats*<sup>-/-</sup>, 1.3 ± 0.3 g; *lats*<sup>+/-</sup>, 1.1 ± 0.1 g). All of these *lats*<sup>-/-</sup> mice were growth retarded. Most of the surviving *lats* homozygous mutant animals gained weight slowly, attaining only 60-70% of normal weight as adults (Figures 7A and B). A representative growth curve for *lats* deficient mice is shown in Figure 7B. To determine if there was a correlation between weight and growth, skeletal growth and organ weight was examined. By radiography, we observed differences in the skull and longitudinal bones that corresponded to the decrease in the size of the mouse. The decreased body weight was not due to the animals being leaner than their wild-type siblings, because the weight of most organs tested had decreased in proportion to the whole body weight. Exceptions were seen in a few particularly reduced organs, including the testis and the ovary.

### Pituitary dysfunction

Male *lats*<sup>-/-</sup> mice displayed decreased fertility although histopathological examination of the testis did not reveal obvious structural abnormalities. *Lats*<sup>-/-</sup> females all displayed severe fertility defects, and approximately 60% of the females were completely sterile. Ovaries from all *lats* deficient females examined contained far fewer follicles than age and parity matched ovaries from *lats*<sup>+/+</sup> females (Figures 8A-D). The majority of follicles observed were primary and secondary follicles. Formation of the antrum was much less prominent than in normal mice. The follicles also contained fewer degenerative granulosa cells, which are common in atretic follicles in normal mice. Moreover, the formation of a corpus luteum was not detected (Figures 8A-D). These histological findings suggest an impairment of ovulation in *lats*<sup>-/-</sup> ovaries, and potential endocrine dysfunction.



The amount of breast epithelial tissue was markedly decreased in *lats*<sup>-/-</sup> females, with some females displaying a complete lack of macroscopic nipple-formation (Figures 9D and E). Histologically, the mammary glands of *lats*<sup>-/-</sup> mice were frequently reduced to a “fat pad” devoid of a ductular component (Figure 9F). This too suggested an endocrine component to the phenotype, as loss of *lats* may alter the levels of hormones that affect breast development.

Estrus is another indicator of endocrine function. Vaginal smears taken from control (+/+) mice showed that they cycled through proestrus, estrus, metestrus, and diestrus in 4 days as described previously for normal mice (Nelson et al., 1982, Biol. Reprod. 27:327-339). In contrast, infertile *lats*<sup>-/-</sup> females did not cycle, and remained in continuous metestrus, an observation that further characterizes their infertility. The abnormal estrus cycle might reflect an underlying problem in signaling between the pituitary and the ovary. To determine if *lats*<sup>-/-</sup> ovaries could respond to appropriate gonadotropin stimulation, young adult mice (7 *lats*<sup>-/-</sup> females; 2 control females) were injected with a Follicle-Stimulating Hormone (FSH) analog (pregnant mare serum), and 46 hours later with a Luteinizing Hormone (LH) preparation (human chorionic gonadotropin). This treatment allowed *lats*<sup>-/-</sup> females to cycle into a prolonged period of estrus, confirming that the temporal synchronization of the levels of endogenous FSH and/or LH in *lats*<sup>-/-</sup> females is deficient.

In all *lats* knock-out mouse pituitaries examined (n=20), hyperplastic changes were readily apparent. There were multiple foci of atypical cells showing irregularly shaped nuclei with an increased content of chromatin and variability in size (Figures 10A and B). This histological atypia was accompanied by a pathological dysfunction of the pituitary. For example, hormone measurements in sera from *lats*<sup>-/-</sup> females revealed: the levels of LH were 3-fold lower than controls (Figure 10C); the PRL levels were less than half of controls (Figure 10D); and Growth Hormone (GH) levels were reduced by 25%. These deficiencies may have resulted from an unbalanced increase in the number of certain types of cells in this normally highly organized tissue. The reduced serum GH level may contribute to the reduced size of *lats*<sup>-/-</sup> mice. The diminished levels of serum LH could account for the lack of proper follicular maturation and differentiation, as well as the infertility observed in female *lats*<sup>-/-</sup> animals, with greater atypia in the pituitary leading to the more severe phenotype. The PRL and LH defects, together, account for both the lack of mammary gland development and the corpus luteum insufficiency syndrome which these animals display.

Interestingly, serum levels of pituitary Follicle Stimulating Hormone (FSH) (Figure 10E) and Thyroid Stimulating Hormone (TSH) are normal. The observation that *lats*<sup>-/-</sup> mice display a selective LH deficiency while sustaining normal FSH production supports a model of differential regulation of the two gonadotropins. This is in accordance with data from Lee and coworkers demonstrating that the transcription factor NGFI-A specifically regulates LH- $\beta$  (Lee et al., 1996, Science 273:1219-1221). *Lats*<sup>-/-</sup> mice thereby provide a model for the human reproductive dysfunction of isolated LH hypogonadotropic hypogonadism.

It is interesting to note that the pituitary deficiencies of *lats*<sup>-/-</sup> mice resembles those of other cell cycle regulator knock-out mice, such as the *Rb*<sup>+/-</sup>, *p53*<sup>-/-</sup>, and *p27*<sup>-/-</sup> mice. In these examples, pituitary cells and other endocrine organs appear to be crucially dependent on cell cycle regulation for their proper development. Indeed, tumor suppressors may play such a key role in the pituitary because critical function in this tissue allows for a link between control of single cell proliferation and total organismal growth and survival.

#### **Tumor development in *lats*<sup>-/-</sup> mice**

Another cause for the infertility of *lats*-deficient females is that they developed ovarian stromal cell tumors by the age of 3 months. The body of the normal ovary consists of spindle-shaped cells, reticular fibers and ground substance which together constitute the ovarian stroma, in which numerous follicles are embedded (Wheater et al., 1987, Functional Histology; Figure 8A). Stromal cell tumors in *lats*<sup>-/-</sup> mice obliterated the normal structure of the ovary, eliminating follicles progressively (Figures 8B and D). These stromal cell tumors are probably not resultant from pituitary dysfunction, as stromal cell tumors are most often local events (Clement, "Histology of the Ovary" in Histology For Pathologists, Second Ed., Sternberg, ed. (Lippincott-Raven, 1997) pp 934-935). Some *lats*<sup>-/-</sup> females were able to give birth to one litter, then became infertile as the stromal cell tumors expanded into the remaining functional ovary. To date, these stromal cell tumors have not yet displayed signs of malignancy. Stromal cell tumors were observed in all *lats* deficient mice examined (n=22), and extended throughout the entire ovary as determined by serial sectioning.

Additional evidence supporting the role of *lats* in mammalian tumorigenesis is that thus far, approximately 15% of all *lats*<sup>-/-</sup> females (n=28) between 4 to 10 months of age have developed large soft tissue sarcomas with metastases to vital organs (e.g. the lungs) (Figures 11A-C). Taking genetic backgrounds into consideration and using litter mates as controls, approximately 57% of all *lats*<sup>-/-</sup> animals developed soft tissue sarcomas. Histology revealed that these sarcomas consist of pleiomorphic, spindle-shaped cells with mitotic figures (Figure 11C), and they displayed immunohistochemical features of fibrosarcomas. No tumors were detected in control mice (n=80). It is possible that the spontaneous rate of tumor formation is even higher in *lats*<sup>-/-</sup> mice as some tumors may be occult, and therefore not readily identifiable. *Lats*<sup>+/-</sup> mice remained tumor free up to 8 months of age. This phenotype is consistent with that observed in heterozygotes for the CDK inhibitors, p27 and p16 (Fero et al., 1996, Cell 85:733-744; Kiyokawa et al., 1996, Cell 85:721-732; Nakayama et al., 1996, Cell 85:707-720; and Serrano et al., 1996, Cell 85:27-37).

The susceptibility of *lats*<sup>-/-</sup> mice to tumor induction by carcinogens was assessed using a standard two-stage protocol consisting of a single application of 9,10-dimethyl-1,2-benzanthracene (DMBA) followed by repeated exposure to ultraviolet B rays (Serrano et al., 1996, Cell 85:27-37). By 7 weeks of age, over 71% of *lats*<sup>-/-</sup> animals developed soft tissue sarcomas, whereas all of the control animals remained free of obvious tumors. The

frequency of tumor development in these *lats*<sup>-/-</sup> animals is particularly impressive, given the fact that C57BL/6 mice are poorly susceptible to the development of skin tumors (as reviewed by DiGiovanni, 1991, Pharm. & Ther. 54:63-128). The spontaneous and induced tumors observed in *lats*<sup>-/-</sup> animals provide clear evidence of the role of *lats* in mammalian tumorigenesis, and attest to the functional conservation of *lats*.

Although both *Drosophila* and mice develop tumors when *lats* is inactivated, the correlation between genotype and phenotype differs between the two organisms. While every *lats*<sup>-/-</sup> cell overproliferates in mosaic flies (Xu et al. 1995, Development 121:1053-1063), only certain tissues in *lats*<sup>-/-</sup> mice develop tumors. This phenotypic difference could be due to the increased complexity in cell cycle control in mammals. For example, in *Drosophila*, *cdc2*/Cyclin A functions at both the G1/S and G2/M transitions in the cell cycle (Whitfield et al., 1990, EMBO J. 9:2563-2573; Knoblich and Lehner, 1993, EMBO J. 12:65-74; Knoblich et al., 1994, Cell 77:117-120; Sauer et al., 1995, Genes & Devel. 9:1237-1239; Sigrist and Lehner, 1997, Cell 90:671-681; Sprenger et al., 1997, Curr. Biol. 7:488-499; Thomas et al., 1997, Genes & Dev. 11:1289-1298). In mammals, however, a different CDK, CDK2, complexes with cyclin A to regulate the G1/S transition, while *cdc2*/cyclin A functions in the G2/M transition (Girard et al., 1991, Cell 67:1169-1179; Tsai et al., 1991, Nature 353:174-177; Hamaguchi et al., 1992, J. Cell. Biol. 117:1041-1053; Pagano et al., 1992, EMBO J. 11:961-967; Rosenblatt et al., 1992, Proc. Natl. Acad. Sci. USA 89:2824-2828; Hunter and Pines, 1994, Cell 79:573-582; Resnitzky et al., 1995, Mol. Cell Biol. 15:4347-4352). We have shown that *lats* negatively regulates *cdc2*/cyclin A, but does not appear to interact with CDK2. Thus, while inactivation of *lats* in *Drosophila* affects regulation of both the G1/S and G2/M transitions of the cell cycle, the removal of *lats* function in mammalian cells only affects control of the G2/M transition. This result is consistent with the fact that human cancers are often caused by mutations in multiple, non-homologous cell cycle control pathways (Kinzler and Vogelstein, 1996, Cell 87:159-170; Weinberg, 1996, Sci. Am. 275:62-70). The higher degree of redundancy in mammalian genomes could also contribute to the phenotypic difference between fly and mouse *lats* mutants.

Although most human tumor suppressors that have been characterized function in the regulation of G1/S (Brown, 1997, Adv. Genet. 36:45-135), there are indications that deregulation of G2/M and M controls may also contribute to tumorigenesis in mammals. For example, p53 is involved in the regulation of both G1/S and G2/M (and M) (Cross et al., 1995, Science 267:1353-1356; Hermeking et al., 1997, Mol. Cell 1:3-11). Inactivation of p53 in mice leads to impressive tumor development, while the disruption of p21, the p53-downstream effector for G1/S, has no tumorigenic effect (Deng et al., 1995, Cell 82:675-684). Recent data also suggests that Rb plays a role in G2/M in addition to its G1/S involvement (Niculescu et al., 1998, Mol. Cell Biol. 18:629-643). Furthermore, overexpression of *cdc2* and mitotic cyclins has been reported in multiple types of human tumors (Arany et al., 1994, Surg. Oncol. 3:153-159; Keyomarsi and Pardee, 1993, Proc.

Natl. Acad. Sci USA 90:1112-1116; Wang et al., 1990, Nature 343:555-557). Signaling through the RHAMM receptor affects cell proliferation by down-regulation of cdc2 and cyclin B transcripts and proteins and results in reversal of tumorigenesis (Mohapatra et al., 1996, J. Exper. Med. 183:1663-1668). The tumorigenic phenotype of *lats*<sup>-/-</sup> mice further supports the notion that the regulation of G2/M or M also plays a role in mammalian tumor development.

A comparison between *lats*- and *p16*-knock-out mice is interesting in several respects. Both *lats* and *p16* are negative regulators of CDKs. In addition, *lats*-knock-out mice resemble *p16*-knock-out animals (Serrano et al., 1996, Cell 85:27-37) in that homozygotes develop tumors at an early age while heterozygotes do not. Although different types of tumors are observed in these two mutants (e.g., ovarian tumors in *lats*<sup>-/-</sup> mice and lymphomas in *p16*<sup>-/-</sup> mice), both types of knock-out mice develop soft tissue sarcomas. The frequencies of spontaneous soft tissue sarcomas between these two knock-out mutants cannot be directly compared due to protocol differences, however, induced tumors in these animals were obtained using the same induction protocol (Serrano et al., 1996, Cell 85:27-37). Interestingly, the frequency of induced tumor formation in *lats*-knock-out mice is even higher than that observed for *p16*<sup>-/-</sup> mice. Over 71% of *lats*<sup>-/-</sup> animals developed soft tissue sarcomas by 7 weeks of age. Only 10% of the *p16*<sup>-/-</sup> mice developed tumors by 7 weeks of age, and 60% of them displayed tumors between the ages of 6-12 weeks (Serrano et al., 1996, Cell 85:27-37).

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Various references are cited herein, the disclosures of which are incorporated by reference in their entireties.

## WHAT IS CLAIMED IS:

1. A recombinant non-human animal in which a *lats* gene has been inactivated by a method comprising introducing a nucleic acid into the animal, or an ancestor thereof, which nucleic acid comprises a non-*lats* sequence flanked by *lats* genomic sequences that promote homologous recombination, such that said non-*lats* sequence replaces the nucleotide sequence encoding the Lats C-terminal domain 1, the Lats C-terminal domain 2, the Lats C-terminal domain 3, and a portion of the kinase domain of the *lats* protein encoded by the *lats* gene.

2. The recombinant non-human animal of claim 1 in which the non-*lats* sequence replaces the nucleotide sequence encoding the amino acids that correspond to amino acids 756-1130 of human *lats*, as depicted in Figure 12 (SEQ ID NO:2).

3. The recombinant non-human animal of claim 1 which is a mouse.

4. The recombinant non-human animal of claim 3 in which the *lats* gene contains a *lats* coding sequence of SEQ ID NO:3.

5. The recombinant non-human animal of claim 1 in which both alleles of the *lats* gene have been inactivated.

6. A method for screening a potential therapeutic compound for activity in treating or preventing cancer comprising administering the compound to the recombinant non-human animal of claim 1; and comparing the size or progression of the cancer in the recombinant non-human animal to which the compound was administered with the size or progression of the cancer in the same recombinant non-human animal prior to administration of the compound or in a recombinant non-human animal that was not so administered or to a standard size or progression of the cancer for such same or a recombinant non-human animal that was not so administered, wherein a decrease in the size or progression of the cancer in the recombinant non-human animal administered the compound as compared to the same animal prior to the administration or to the recombinant non-human animal not so administered or to the standard size or progression of the cancer, indicates that the compound has activity in treating or preventing cancer.

7. The method of claim 6 in which the non-*lats* sequence replaces the nucleotide sequence encoding the amino acids that correspond to amino acids 756-1130 of human *lats*, as depicted in Figure 12 (SEQ ID NO:2).

8. The method of claim 6 in which the recombinant non-human animal is a mouse.

9. The method of claim 8 in which the *lats* gene contains the *lats* coding sequence of SEQ ID NO:3.

10. The method of claim 6 in which both alleles of the *lats* gene have been inactivated.

11. The method of claim 6 in which the compound is screened for activity in treating or preventing soft tissue sarcomas.

12. The method of claim 6 in which the compound is screened for activity in treating or preventing ovarian tumors.

13. A method for screening a potential therapeutic compound for activity in treating or preventing cancer comprising recombinantly expressing the compound in the recombinant non-human animal of claim 1; and comparing the size or progression of the cancer in the recombinant non-human animal in which the compound was expressed with the size or progression of the cancer in the same recombinant non-human animal prior to expression of the compound or in a recombinant non-human animal in which the compound was not so expressed or to a standard size or progression of the cancer for such same or a recombinant non-human animal in which the compound was not so expressed, wherein a decrease in the size or progression of the cancer in the recombinant non-human animal in which the compound was expressed as compared to the same animal prior to the expression of the compound or to the recombinant non-human animal in which said compound was not so expressed or to the standard size or progression of the cancer, indicates that the compound has activity in treating or preventing cancer.

14. A method for screening a potential therapeutic compound for activity in treating or preventing skin cancer comprising administering the compound to a *lats* knock-out animal having skin tumors induced by exposure to at least one carcinogen; and comparing the size or progression of the skin tumors on the *lats* knock-out animal to which the compound was administered with the size or progression of skin cancers on the same *lats* knock-out animal prior to administration of the compound or on a *lats* knock-out animal in which skin tumors have also been induced by exposure to said at least one carcinogen but which has not been administered the compound or to a standard size or progression of the skin tumors for such same or a *lats* knock-out animal that was not so administered, wherein a reduction in the size or progression of the skin tumors in the *lats* knock-out animal administered the compound as compared to the same animal prior to

administration of the compound or to the animal not so administered or to the standard size or progression of the skin tumors, indicates that the compound has activity in treating or preventing skin cancer.

5        15.     The method of claim 14 in which the *lats* knock-out animal has at least one *lats* gene which was inactivated by promoting homologous recombination between *lats* genomic sequences and a nucleic acid having non-*lats* sequences flanked by genomic sequences.

10       16.     The method of claim 15 in which the non-*lats* sequence replaces the nucleotide sequence encoding the Lats C-terminal domain 1, the Lats C-terminal domain 2, the Lats C-terminal domain 3, and a portion of the kinase domain of the *lats* protein encoded by the *lats* gene.

15       17.     The method of claim 15 in which the non-*lats* sequence replaces the nucleotide sequence encoding the amino acids that correspond to amino acids 756-1130 of human *lats*, as depicted in Figure 12 (SEQ ID NO:2).

18.     The method of claim 14 in which the *lats* knock-out animal is a mouse.

20       19.     The method of claim 18 in which the *lats* gene contains the *lats* coding sequence of SEQ ID NO:3.

20.     The method of claim 14 in which both alleles of the *lats* gene have been inactivated.

25       21.     The method of claim 14 in which the skin tumors were induced by 9,10-dimethyl-1,2-benzanthracene and repeated exposure to ultraviolet B radiation.

22.     The method of claim 14 in which the potential therapeutic compound is administered topically.

30       23.     A method for screening a potential therapeutic compound for activity in treating or preventing skin cancer comprising recombinantly expressing the compound in a *lats* knock-out animal having skin tumors induced by exposure to at least one carcinogen; and comparing the size or progression of the skin tumors on the *lats* knock-out animal in which the compound was expressed with the size or progression of skin cancers on the same *lats* knock-out animal prior to expression of the compound or on a *lats* knock-out animal in  
35       which skin tumors have also been induced by exposure to said at least one carcinogen but in which the compound has not been expressed or to a standard size or progression of the skin

tumors for such same or a *lats* knock-out animal in which the compound was not so expressed, wherein a reduction in the size or progression of the skin tumors in the *lats* knock-out animal in which the compound was expressed as compared to the same animal prior to expression of the compound or to the animal in which the compound was not so expressed or to the standard size or progression of the skin tumors, indicates that the compound has activity in treating or preventing skin cancer.

24. A method for screening a potential therapeutic compound for activity in treating or preventing a disease or disorder associated with pituitary dysfunction comprising administering the compound to a *lats* knock-out animal; and comparing the level of an indicator of pituitary dysfunction in the *lats* knock-out animal to which the compound has been administered to the level of the indicator in the same *lats* knock-out animal prior to administration of the compound or to a *lats* knock-out animal that has not been administered the compound or to a standard level of the indicator for such same or a *lats* knock-out animal that was not so administered, wherein a change in the indicator toward the level of the indicator in a wild type animal not afflicted with a pituitary dysfunction as compared to the same animal prior to administration of the compound or to the animal not so administered or to the standard level of the indicator, indicates that the compound is active to treat or prevent a disease or disorder associated with pituitary dysfunction.

25. The method of claim 24 in which the *lats* knock-out animal has at least one *lats* gene which was inactivated by promoting homologous recombination between *lats* genomic sequences and a nucleic acid having non-*lats* sequences flanked by genomic sequences.

26. The method of claim 25 in which the non-*lats* sequence replaces the nucleotide sequence encoding the Lats C-terminal domain 1, the Lats C-terminal domain 2, the Lats C-terminal domain 3, and a portion of the kinase domain of the *lats* protein encoded by the *lats* gene.

27. The method of claim 25 in which the non-*lats* sequence replaces the nucleotide sequence encoding the amino acids that correspond to amino acids 756-1130 of human *lats*, as depicted in Figure 12 (SEQ ID NO:2).

28. The method of claim 24 in which the *lats* knock-out animal is a mouse.

29. The method of claim 28 in which the *lats* gene contains a *lats* coding sequence of SEQ ID NO:3.



30. The method of claim 24 in which both alleles of the *lats* gene have been inactivated.

31. The method of claim 24 in which the indicator is fertility.

32. The method of claim 24 in which the indicator is ovulation.

33. The method of claim 24 in which the indicator is linear growth.

34. The method of claim 24 in which the indicator is serum levels of luteinizing hormone, growth hormone or prolactin.

35. The method of claim 24 in which the disease or disorder is LH hypogonadotropic hypogonadism.

36. A method for screening a potential therapeutic compound for activity in treating or preventing a disease or disorder associated with pituitary dysfunction comprising recombinantly expressing the compound in a *lats* knock-out animal; and comparing the level of an indicator of pituitary dysfunction in the *lats* knock-out animal in which the compound has been expressed to the level of the indicator either in the same *lats* knock-out animal prior to expression of the compound or to a *lats* knock-out animal in which the compound has not been expressed or to a standard level of the indicator for such same or a *lats* knock-out animal in which the compound was not so expressed, wherein a change in the indicator toward the level of the indicator in a wild type animal not afflicted with a pituitary dysfunction as compared to the same animal prior to expression of the compound or to the animal in which the compound was not so expressed or to the standard level of the indicator, indicates that the compound is active to treat or prevent a disease or disorder associated with pituitary dysfunction.

37. The method of claim 6, 14 or 24 in which the compound is purified.

38. A method for treating a cancer that has been shown to be refractory to a chemotherapy or radiation therapy in a subject in need of such treatment comprising administering to the subject a therapeutically effective amount of a molecule that promotes *lats* function.

39. The method of claim 38 in which the subject is a human.

40. The method of claim 38 in which the molecule is a *lats* protein.

41. The method of claim 38 in which the molecule is a protein having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, and SEQ ID NO:8.

5 42. The method of claim 38 in which the molecule is a protein having the amino acid sequence of SEQ ID NO:2.

43. The method of claim 38 in which the molecule is a lats analog or derivative that has activity to promote lats function.

10 44. The method of claim 38 in which the molecule is a protein encoded by a first nucleic acid that is hybridizable under conditions of low stringency to a second nucleic acid having a nucleotide sequence that is the reverse complement of a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, and SEQ ID NO:7, said protein having activity to inhibit cell overproliferation.

15 45. The method of claim 38 in which the molecule is a protein consisting of at least 20 contiguous amino acids of a protein having an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, and SEQ ID NO:8, said protein having activity to inhibit cell overproliferation.

20 46. The method of claim 38 in which the molecule is a protein comprising a domain of a lats protein selected from the group consisting of a lats C-terminal domain 3 (LCD3), lats C-terminal domain 2 (LCD2), lats C-terminal domain 1 (LCD1), kinase domain, lats flanking domain (LFD), lats split domain 1 (LSD1), lats split domain 2 (LSD2), and SH3-binding domain, said protein having activity to inhibit cell  
25 overproliferation.

47. The method of claim 40 in which the lats protein is phosphorylated.

48. The method of claim 47 in which the lats protein is phosphorylated on a serine or threonine residue within 20 residues upstream of the amino acid sequence Ala-Pro-  
30 Glu in a subdomain eight of a kinase domain of said lats protein.

49. The method of claim 48 in which the lats protein is phosphorylated at a serine residue corresponding to serine 909 of the human lats amino acid sequence, as depicted in Figure 12 (SEQ ID NO:2).

35 50. The method of claim 43 in which the lats analog or derivative has a threonine or serine residue within 20 residues upstream of the amino acid sequence Ala-Pro-Glu in a

subdomain eight of a kinase domain of said lats analog or derivative substituted with an aspartate or glutamate residue.

51. The method of claim 50 in which the lats analog or derivative has a glutamate residue substituted for a serine residue at the residue corresponding to serine 909 of the human lats amino acid sequence, as depicted in Figure 12 (SEQ ID NO:2).

52. The method of claim 38 in which said molecule is a chimeric protein comprising a fragment of a lats protein, said fragment consisting of at least 20 contiguous amino acids of said lats protein, fused via a covalent bond to an amino acid sequence of a second protein, said second protein not being a lats protein, said chimeric protein having activity to inhibit cell overproliferation.

53. The method of claim 38 in which said cancer has been shown to be refractory to radiation therapy.

54. The method of claim 38 in which said cancer has been shown to be refractory to chemotherapy.

55. The method of claim 54 in which said chemotherapy kills cancer cells during S phase of the cell cycle.

56. The method of claim 54 in which said chemotherapy kills cancer cells during mitosis.

57. The method of claim 38 which further comprises administering one or more chemotherapeutic agents to the subject.

58. The method of claim 57 in which said one or more chemotherapeutic agents are administered concurrently with the administration of said molecule.

59. The method of claim 57 in which said one or more chemotherapeutic agents are administered subsequent to the administration of said molecule.

60. The method of claim 38 in which said molecule is a nucleic acid comprising a nucleotide sequence encoding a lats protein.

61. The method of claim 60 in which said nucleotide sequence is SEQ ID NO:1.

62. The method of claim 60 in which said nucleotide sequence encodes a protein having the amino acid sequence of SEQ ID NO:2.

5 63. The method of claim 38 in which said molecule is a first nucleic acid that hybridizes under low stringency conditions to a second nucleic acid that is the reverse complement of a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, and SEQ ID NO:7.

10 64. A method for treating a cancer that has been shown to be refractory to a chemotherapy or radiation therapy in a subject in need of such treatment comprising administering to the subject a therapeutically effective amount of a cell that expresses a recombinant nucleic acid that promotes lats function.

65. The method of claim 64 in which said nucleic acid comprises the nucleotide sequence of SEQ ID NO:1.

15 66. The method of claim 64 in which said nucleic acid comprises a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO:2.

20 67. A kit comprising in one or more containers a therapeutically effective amount of a molecule selected from the group consisting of a lats protein, a lats derivative, a lats analog, a nucleic acid encoding a lats protein, a nucleic acid encoding a lats derivative, and a nucleic acid encoding a lats analog; and at least one chemotherapeutic agent.

68. A purified complex of a lats protein and a cdc2 protein.

25 69. The purified complex of claim 68 in which the proteins are human proteins.

70. The purified complex of claim 68 in which the lats protein is phosphorylated.

30 71. The purified complex of claim 70 in which the lats protein is phosphorylated on a serine or threonine residue within 20 residues upstream of the amino acid sequence Ala-Pro-Glu in a subdomain eight of a kinase domain of said lats protein.

72. The purified complex of claim 71 in which the lats protein is phosphorylated at a serine residue corresponding to serine 909 of the human lats amino acid sequence, as depicted in Figure 12 (SEQ ID NO:2).

35 73. A purified complex selected from the group consisting of a complex of a derivative of a lats and a cdc2 protein, a complex of a lats protein and a derivative of a cdc2,

and a complex of a derivative of a lats protein and a derivative of a cdc2 protein, in which the derivative of the lats protein is able to form a complex with a wild-type cdc2 protein and the derivative of the cdc2 is able to form a complex with a wild-type lats protein.

5           74.     The purified complex of claim 73 in which the derivative of the lats protein and/or the cdc2 protein is fluorescently labeled.

          75.     The purified complex of claim 73 in which the lats derivative has a threonine or serine residue within 20 residues upstream of the amino acid sequence Ala-Pro-Glu in a subdomain eight of a kinase domain of said lats derivative substituted with an aspartate or  
10    glutamate residue.

          76.     The purified complex of claim 75 in which the lats derivative has a glutamate residue substituted for a serine residue at the residue corresponding to serine 909 of the human lats amino acid sequence, as depicted in Figure 12 (SEQ ID NO:2).

15           77.     The purified complex of claim 73 in which the lats derivative is a fragment of a lats protein consisting of the amino acid sequence corresponding to amino acids 15-585 of the amino acid sequence of human lats, as depicted in Figure 12 (SEQ ID NO:2).

          78.     A chimeric protein comprising a fragment of a lats protein consisting of at  
20    least 6 amino acids fused via a covalent bond to a fragment of a cdc2 protein consisting of at least 6 amino acids.

          79.     The chimeric protein of claim 78 in which the fragment of the lats protein is a fragment capable of binding the cdc2 protein and in which the fragment of the cdc2  
25    protein is a fragment capable of binding the lats protein.

          80.     The chimeric protein of claim 78 in which the fragment of the lats protein has an amino acid sequence corresponding to amino acids 15 to 585 of the amino acid sequence of human lats, as depicted in Figure 12 (SEQ ID NO:2).

30           81.     The chimeric protein of claim 79 in which the fragment of the lats protein and the fragment of the cdc2 protein form a lats-cdc2 complex.

          82.     An antibody which immunospecifically binds the complex of claim 68 or a fragment or derivative of said antibody containing the binding domain thereof.

35           83.     The antibody of claim 82 which does not immunospecifically bind a lats protein or a cdc2 protein that are not part of a lats-cdc2 complex.

84. An isolated nucleic acid or an isolated combination of nucleic acids comprising a nucleotide sequence encoding a lats protein and a nucleotide sequence encoding a cdc2 protein.

5 85. The isolated nucleic acid or isolated combination of nucleic acids of claim 84 which are nucleic acid vectors.

86. An isolated nucleic acid that comprises a nucleotide sequence encoding the chimeric protein of claim 78.

10 87. A cell containing the nucleic acid of claim 84, which nucleic acid is recombinant.

88. A cell containing the nucleic acid of claim 86, which nucleic acid is recombinant.

15 89. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of the complex of claim 68; and a pharmaceutically acceptable carrier.

20 90. The pharmaceutical composition of claim 89 in which the proteins are human proteins.

91. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of the complex of claim 73; and a pharmaceutically acceptable carrier.

25 92. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of the chimeric protein of claim 79; and a pharmaceutically acceptable carrier.

30 93. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of the antibody of claim 83 or a fragment or derivative of said antibody containing the binding domain thereof; and a pharmaceutically acceptable carrier.

35 94. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of the nucleic acids or combination of nucleic acids of claim 84; and a pharmaceutically acceptable carrier.

95. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of the isolated nucleic acid of claim 86; and a pharmaceutically acceptable carrier.

5 96. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of the recombinant cell of claim 87; and a pharmaceutically acceptable carrier.

97. A pharmaceutical composition comprising a therapeutically or prophylactically effective amount of the recombinant cell of claim 88; and a  
10 pharmaceutically acceptable carrier.

98. A method of producing a complex of a lats protein and a cdc2 protein comprising growing a recombinant cell containing the nucleic acid of claim 84 such that the encoded lats and cdc2 proteins are expressed and bind to each other, and recovering the expressed complex of the lats protein and the cdc2 protein.  
15

99. A method of diagnosing or screening for the presence of or a predisposition for developing a disease or disorder characterized by an aberrant level of a complex of a lats protein and a cdc2 protein in a subject comprising measuring the level of said complex, RNA encoding the lats and the cdc2 proteins, or functional activity of said complex, in a  
20 sample derived from the subject, in which an increase or decrease in the level of said complex, said RNA encoding lats and cdc2, or functional activity of said complex in the sample, relative to the level of said complex, said RNA encoding lats and cdc2, or functional activity of said complex found in an analogous sample from a subject not having the disease or disorder or a predisposition for developing the disease or disorder, indicates  
25 the presence of the disease or disorder or a predisposition for developing the disease or disorder.

100. A kit comprising in one or more containers a substance selected from the group consisting of a complex of a lats and a cdc2 protein, an antibody against said complex, nucleic acid probes capable of hybridizing to RNA of lats and RNA of cdc2, or  
30 pairs of nucleic acid primers capable of priming amplification of at least a portion of a gene for lats and a gene for cdc2.

101. A method for modulating the activity of cdc2 comprising administering a molecule that promotes, inhibits, or antagonizes lats function.

35 102. A method for inhibiting the activity of cdc2 comprising administering a molecule that promotes lats function.

103. A method for increasing the activity of cdc2 comprising administering a molecule that inhibits or antagonizes lats function.

5 104. A method for treating or preventing a disease or disorder associated with an aberrantly high level of cdc2 in a subject in need of such treatment or prevention comprising administering to the subject a therapeutically effective amount of a molecule that promotes lats function.

10 105. The method of claim 104 in which said molecule is selected from the group consisting of a lats protein, a lats derivative or analog that promotes lats function, a nucleic acid encoding a lats protein, and nucleic acid encoding a lats derivative or analog that promotes lats function, and a lats agonist.

15 106. A method for treating or preventing a disease or disorder associated with an aberrantly low level of cdc2 activity in a subject in which such treatment or prevention is desired comprising administering to the subject a therapeutically effective amount of a molecule that inhibits or antagonizes lats function.

20 107. The method of claim 106 in which said molecule is selected from the group consisting of a lats analog or derivative that inhibits or antagonizes lats function, an anti-lats antibody, and a *lats* antisense nucleic acid.

25 108. A method for screening a molecule for efficacy in treating or preventing a cancer refractory to chemotherapy or radiation therapy, said method comprising contacting cancer cells that are refractory to treatment with chemotherapeutic agents or radiation with the molecule and comparing the proliferation or survival of the contacted cells with the proliferation or survival of cells not so contacted, wherein a lower level of proliferation or survival of the contacted cells indicates that the molecule is effective to treat or prevent the cancer.

30 109. The method of claim 108 in which said cells are cultured *in vitro* from a tissue sample of a patient.

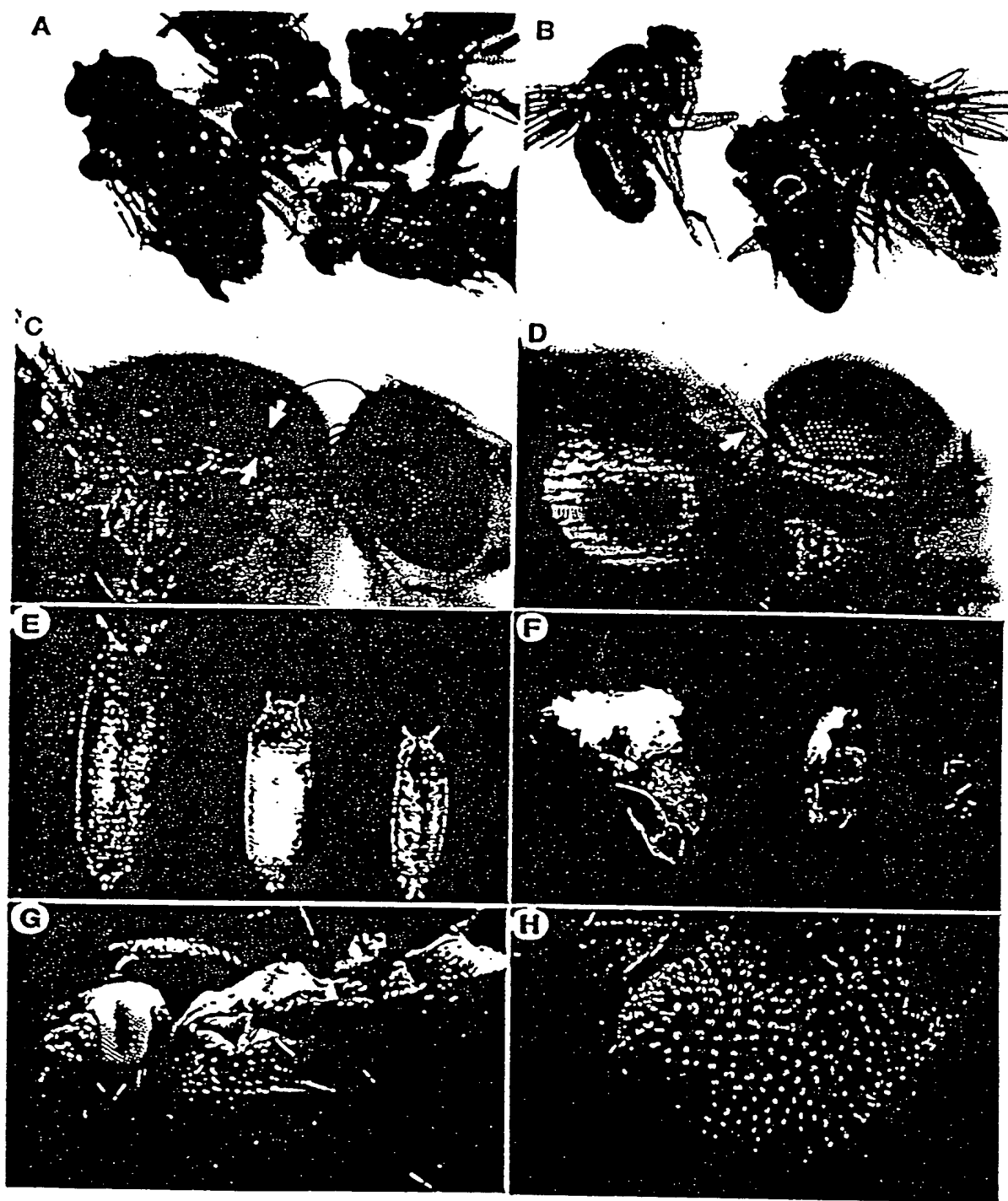
35 110. A method for screening a molecule for activity to modulate cdc2 levels or activity comprising contacting cells with the molecule, and comparing the level of cdc2 protein, mRNA or activity in cells contacted with the molecule to the amount of cdc2 protein, mRNA, or activity in cells not so contacted, wherein an increase or decrease in the amount of cdc2 protein, mRNA, or activity in the contacted cells relative to the amount of cdc2 protein, mRNA, or activity in the cells not so contacted indicates that the molecule has activity to modulate cdc2 levels or activity.



111. A method for screening a molecule for activity to modulate, directly or indirectly, the formation of a complex of lats and cdc2 proteins comprising measuring the levels of said complex formed from lats and cdc2 proteins in the presence of said molecule under conditions conducive to formation of the complex; and comparing the levels of said complex with the levels of said complex that are formed in the absence of said molecule, wherein a lower or higher level of said complex in the presence of said molecule indicates that the molecule modulates formation of said complex.

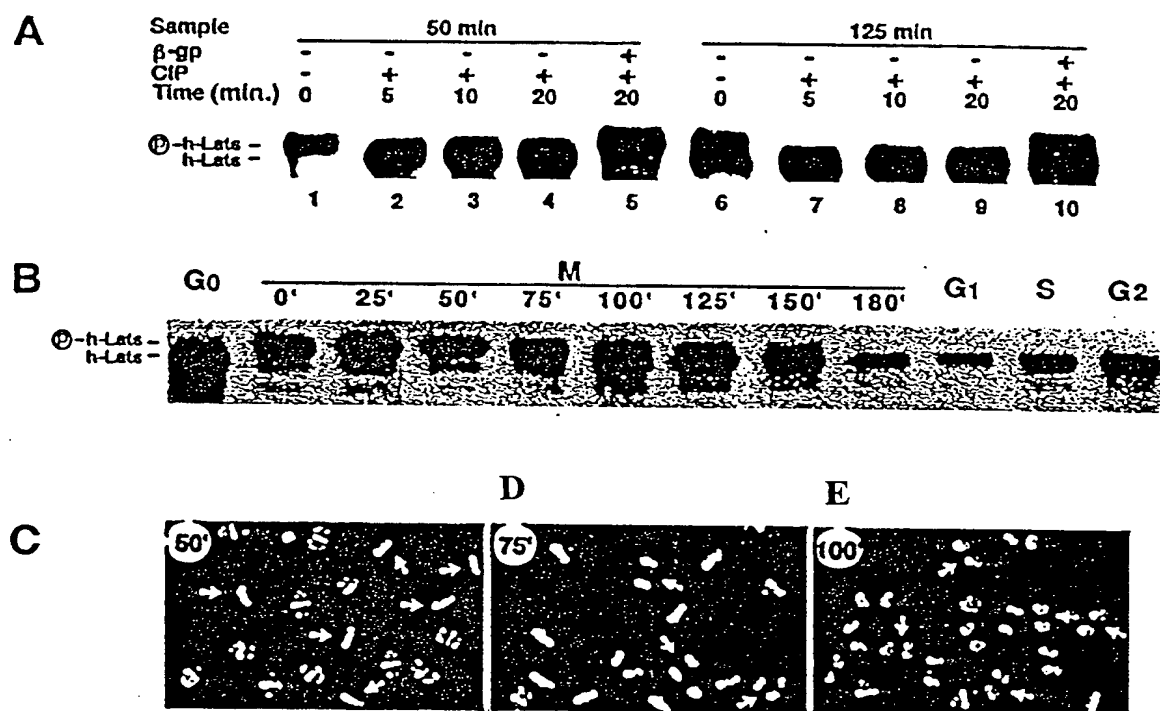
112. The method of claim 111 in which the molecule inhibits formation of the complex.

113. The method of claim 111 in which the molecule promotes formation of the complex.

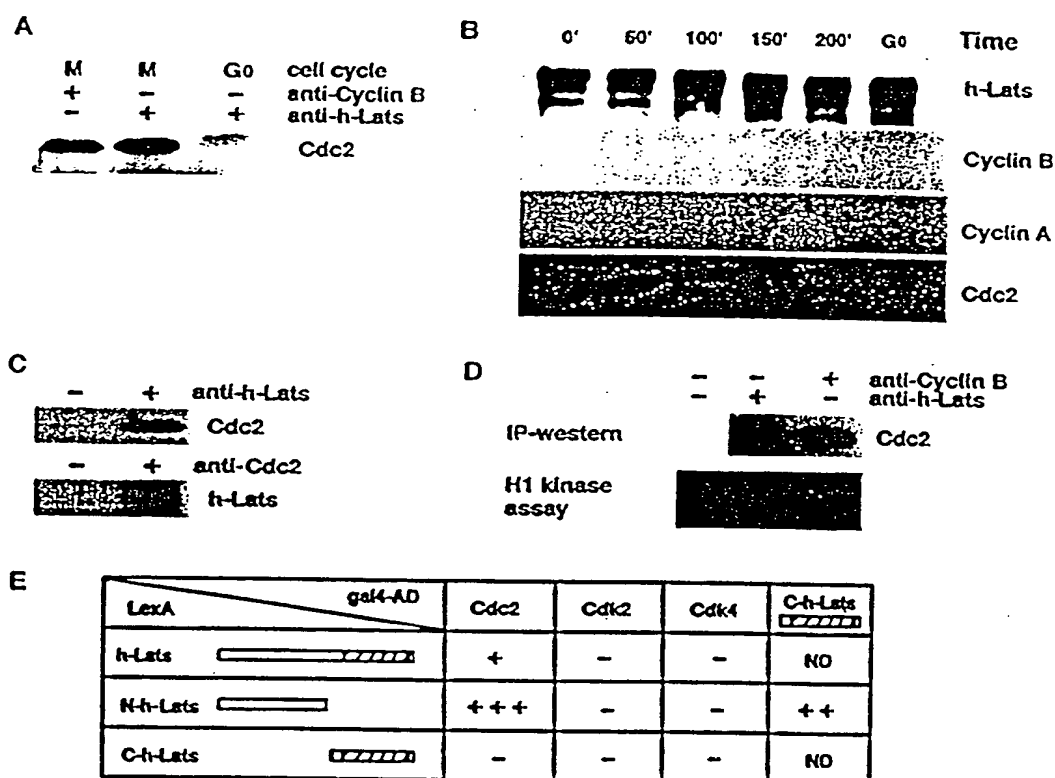


FIGS. 1A-H

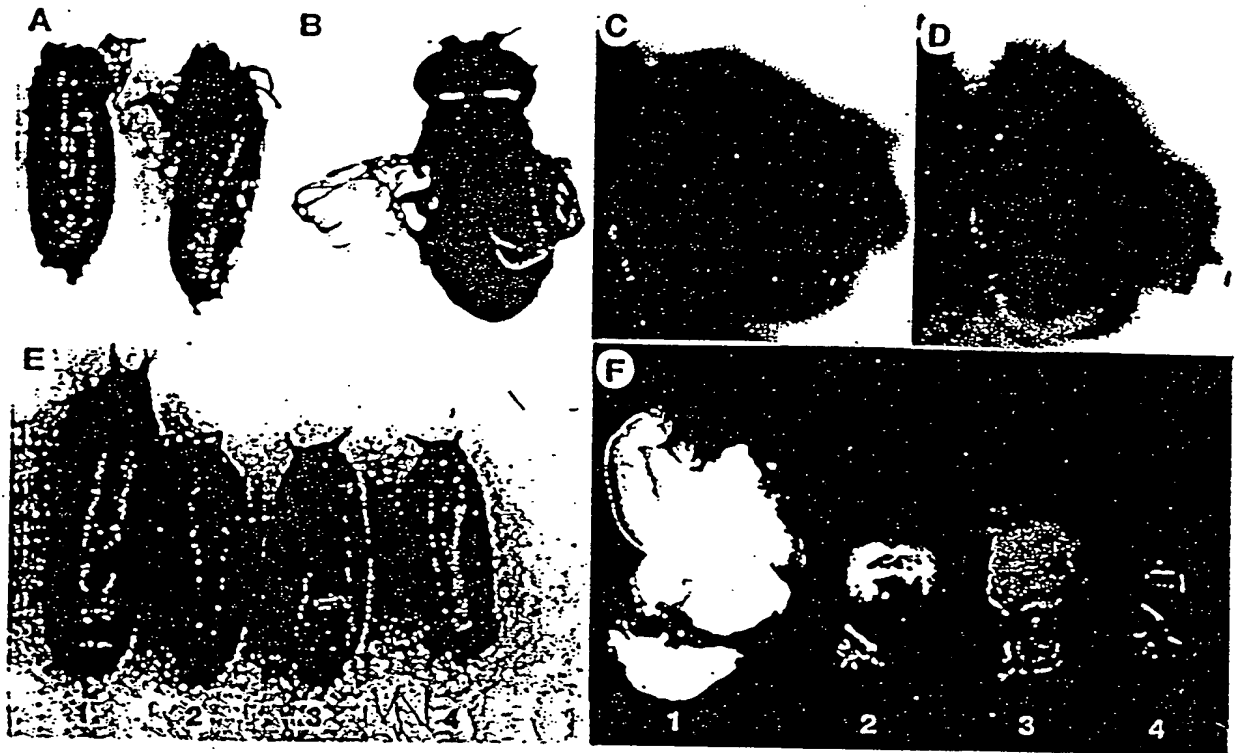
2/25



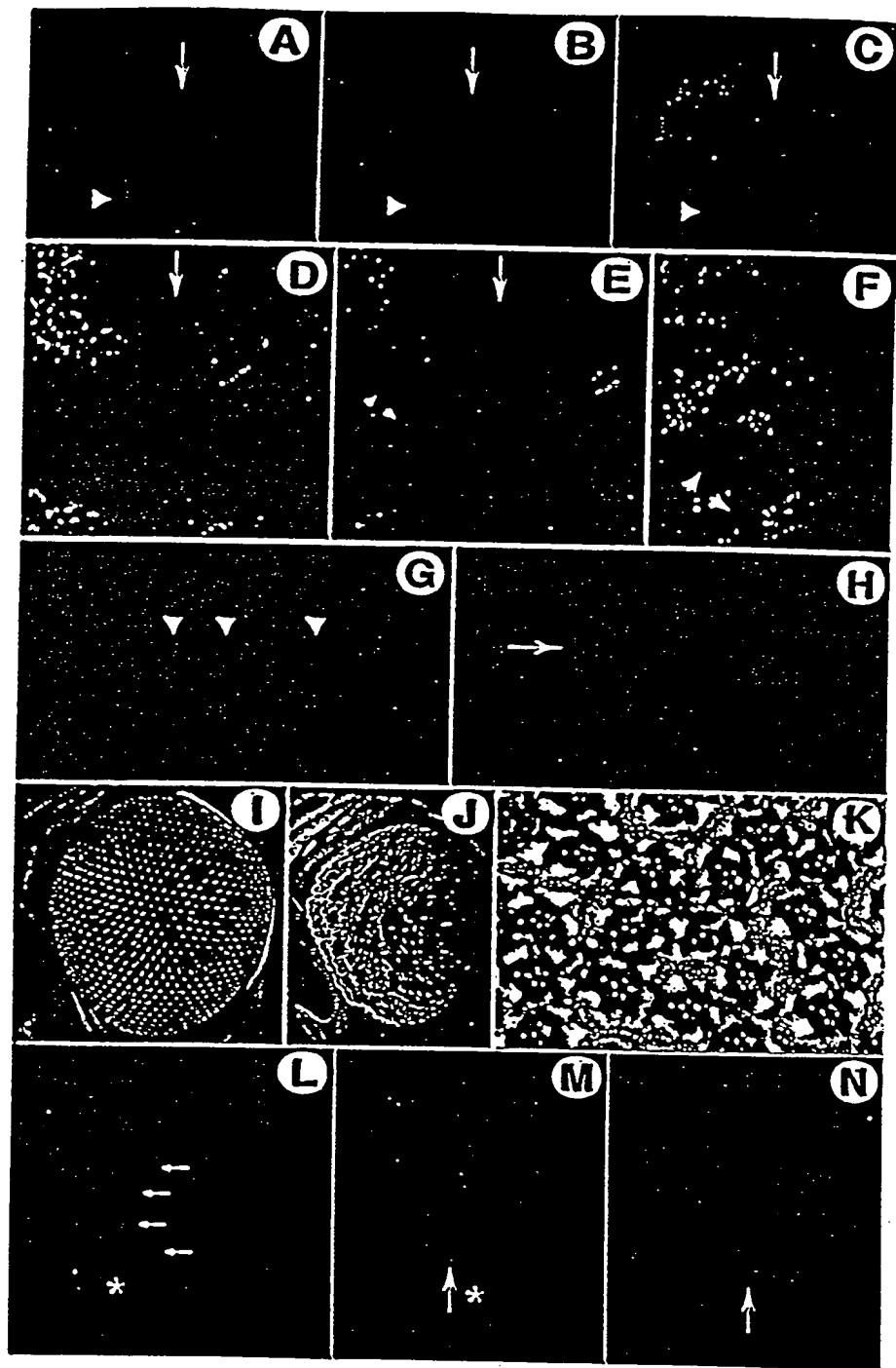
FIGS. 2A-E



FIGS. 3A-E



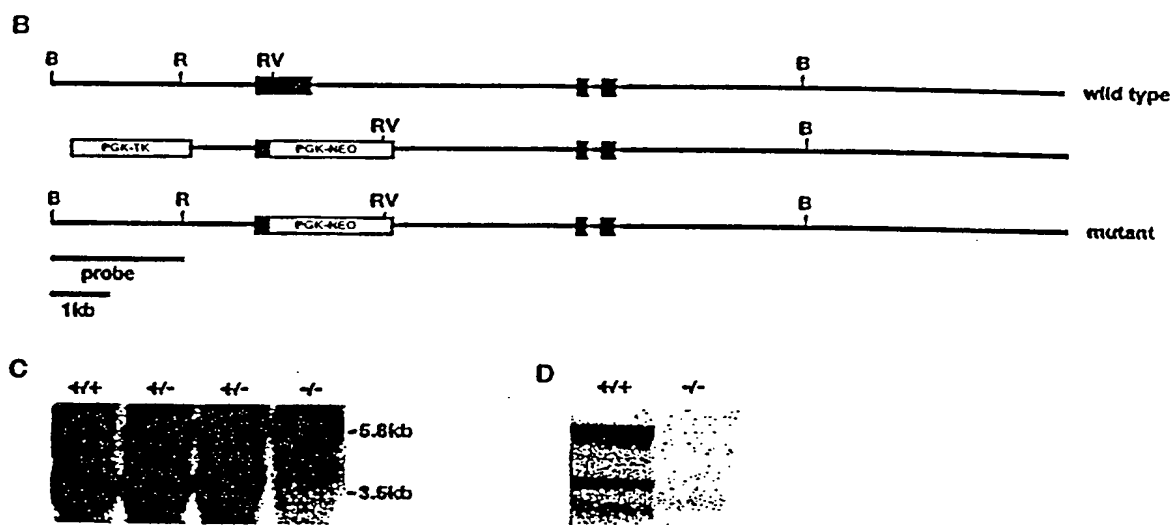
**FIGS. 4A-F**



FIGS. 5A-N

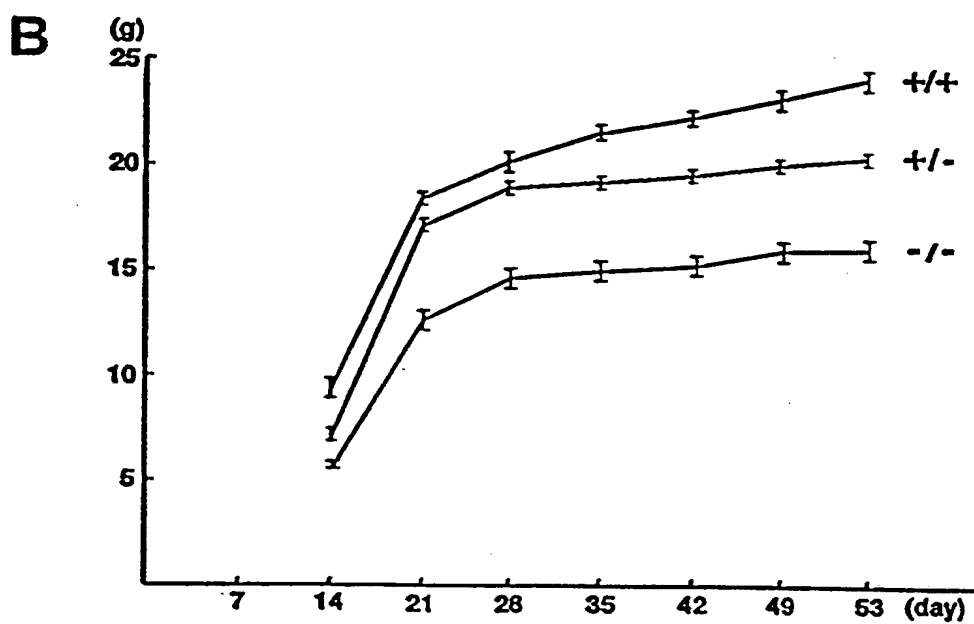
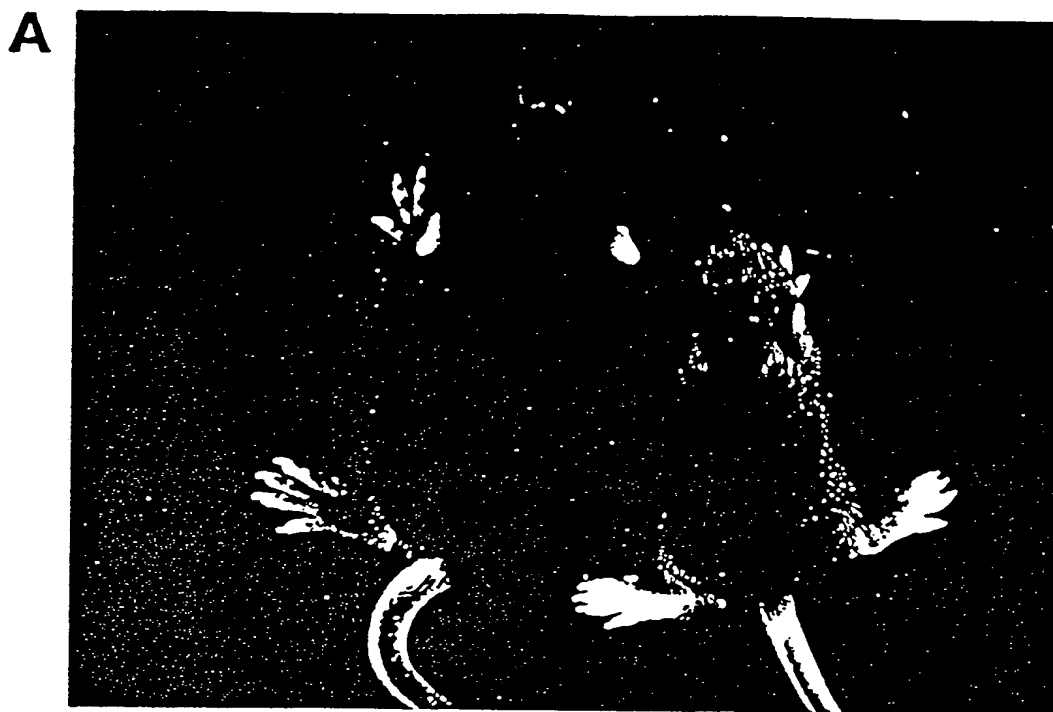
1 MKRSEKPEGYRQMRPKT FPASNYTVSSRQMLQE IRESLRNLKPSDAAKAEHNMSKMSTEDPRQVRNPPK h-Lats  
 71 FGTHHKALQEI RNSL LPFANETNSSRSTSEVNPQMLQDLQAAGFDEDMVIAQLQKTNRRSIEAAIEFISK h-Lats  
 141 MSYQDPRREQMAAAAARPINASMKP GNVVDSYNHKAUSWAGSKESLVPOHHGPPGESVATHSESPNSOIT h-Lats  
 EVOFSTNRKOSWKGSKESSLVPOHHGPPGENVVYRSESPNSOIA h-Lats  
 211 /GRPLSGSGISAEVJAHPSNGORVNPFPPOVHSVIPP PPHGOTPPPHGTIPPPPSWE PINSOTKHYSGN h-Lats  
 /GRPLSGSGIAAEVJAHPSNGORVNPFPPOVRSVTPPPPPRGOTPPPRGTTPPPPSWEPSOTKHYSGN h-Lats  
 281 MEYVISHISPVPPGAVOEGYPPPPPIITSPMNPSSADRAISSVPVGRPIIMOSTSKFNFPSEHPGMJNG h-Lats  
 MEYVISHISPVPPGAVOEGYPPPPPIITSPMNPSSADRAISSVPVGRPIIMOSTSKFNFPSEHPGMJNG h-Lats  
 351 TETIDJMIHUNVVIAITINOPPPPYPLTAANGOSPSALOTGSAAPSSYTNGSIPOSMMVNPNNHSHNME h-Lats  
 GFSODFIVHONV-DTISITOPPPPYPLTPANGOSPSALOTGSAAPSSYFANGVNPPOSMMVNPNNHSHNME h-Lats  
 421 -YNISVPGLOTNPOSSSAPAUSSPSSEHEIPTWOPNIPVRSNSFNPLGNASHSANSOPSATTVIATIT h-Lats  
 YNINVPGLONTNPOSSSAPAUSSPSSEHEIPTWOPNIPVRSNSFNPLGNASHSANSOPSATTVIATIT h-Lats  
 491 PAPIQOPVKSMRVLKPELOALAPTHPSWIPOPVTVOPSPFEGTASSVTPPVAEAPSYOGPPPPYP h-Lats  
 PAPIQOPVKSMRVLKPELOALAPTHPSWIPOPVTVOPSPFEGTASSVTPPVAEAPSYOGPPPPYP h-Lats  
 561 KHLHONPSVPPYESISKPSKEDJPSLPKEDSEKSYENVDSGOKKKOITTSPI TVRKNKKDEERRESR h-Lats  
 KHLHONPSVPPYESISKPSKEDJPSLPKEDSEKSYENVDSGOKKKOITTSPI TVRKNKKDEERRESR h-Lats  
 631 IOSYSPQAFKFFMEOHVENVLKSHOORLHKKOLENEMMRVGLSODAODOMRKMMLCOKE SNYIRLKRKAKM h-Lats  
 IOSYSPQAFKFFMEOHVENVLKSHOORLHKKOLENEMMRVGLSODAODOMRKMMLCOKE SNYIRLKRKAKM h-Lats  
 701 DKSMFVKIKITLGI GAFGEVCLARKVDIKALYATKILHKKDVLLENOVAHVKAERDILAEADNEWVVRLYY h-Lats  
 DKSMFVKIKITLGI GAFGEVCLARKVDIKALYATKILHKKDVLLENOVAHVKAERDILAEADNEWVVRLYY h-Lats  
 771 SFQDKONLYFVMDYIPGGDMSSLIRMGIFPENLAHFYIAELTCAVESVHKMGFIHRDIKPDNITIDRDG h-Lats  
 SFQDKONLYFVMDYIPGGDMSSLIRMGIFPENLAHFYIAELTCAVESVHKMGFIHRDIKPDNITIDRDG h-Lats  
 841 HIKLIDFGLCTGFHWIHD SKY YOSGDHPRDSMDFSNEWGDPSCRCGDRLKPLEHRAAHOHOC LAHSL h-Lats  
 HIKLIDFGLCTGFHWIHD SKY YOSGDHPRDSMDFSNEWGDPSCRCGDRLKPLEHRAAHOHOC LAHSL h-Lats  
 911 VGI PNYIAPEVLLRTGYTOLCDWMSVGVILCEMLVGOPPF LAOTPLETOMKVINWOTSLHIPPOAKLSPE h-Lats  
 VGI PNYIAPEVLLRTGYTOLCDWMSVGVILCEMLVGOPPF LAOTPLETOMKVINWOTSLHIPPOAKLSPE h-Lats  
 981 ASDLIKLCRGPEDR LGKNGADEIKAHFFKTI DFSSDLROOSASYIPKITHPTDTSNFDVPDPK LWSU h-Lats  
 ASDLIKLCRGPEDR LGKNGADEIKAHFFKTI DFSSDLROOSASYIPKITHPTDTSNFDVPDPK LWSU h-Lats  
 1051 DNEENVNDILNSWYKNGKHPERAFYEFTFRHFDDNGYPYNYPKPIEY EYINSQGEQOSDEDDONIGS h-Lats  
 GSEENISDTLSWYKNGKHPERAFYEFTFRHFDDNGYPYNYPKPIEY EYINSQGEQOSDEDDONIGS h-Lats  
 1121 EIKNRDLVYV h-Lats  
 DGNNRDLVYV h-Lats

FIG. 6A



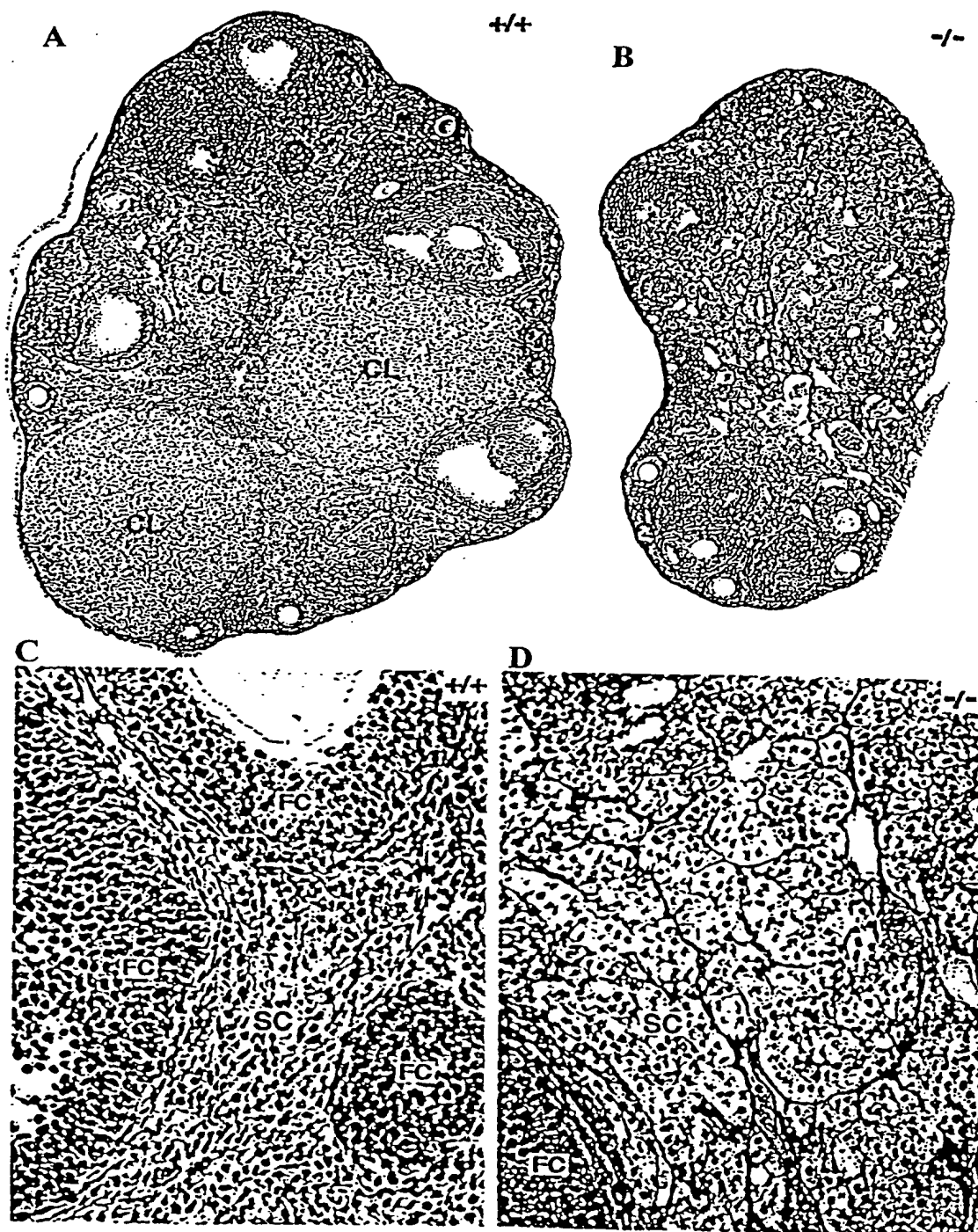
FIGS. 6B-D





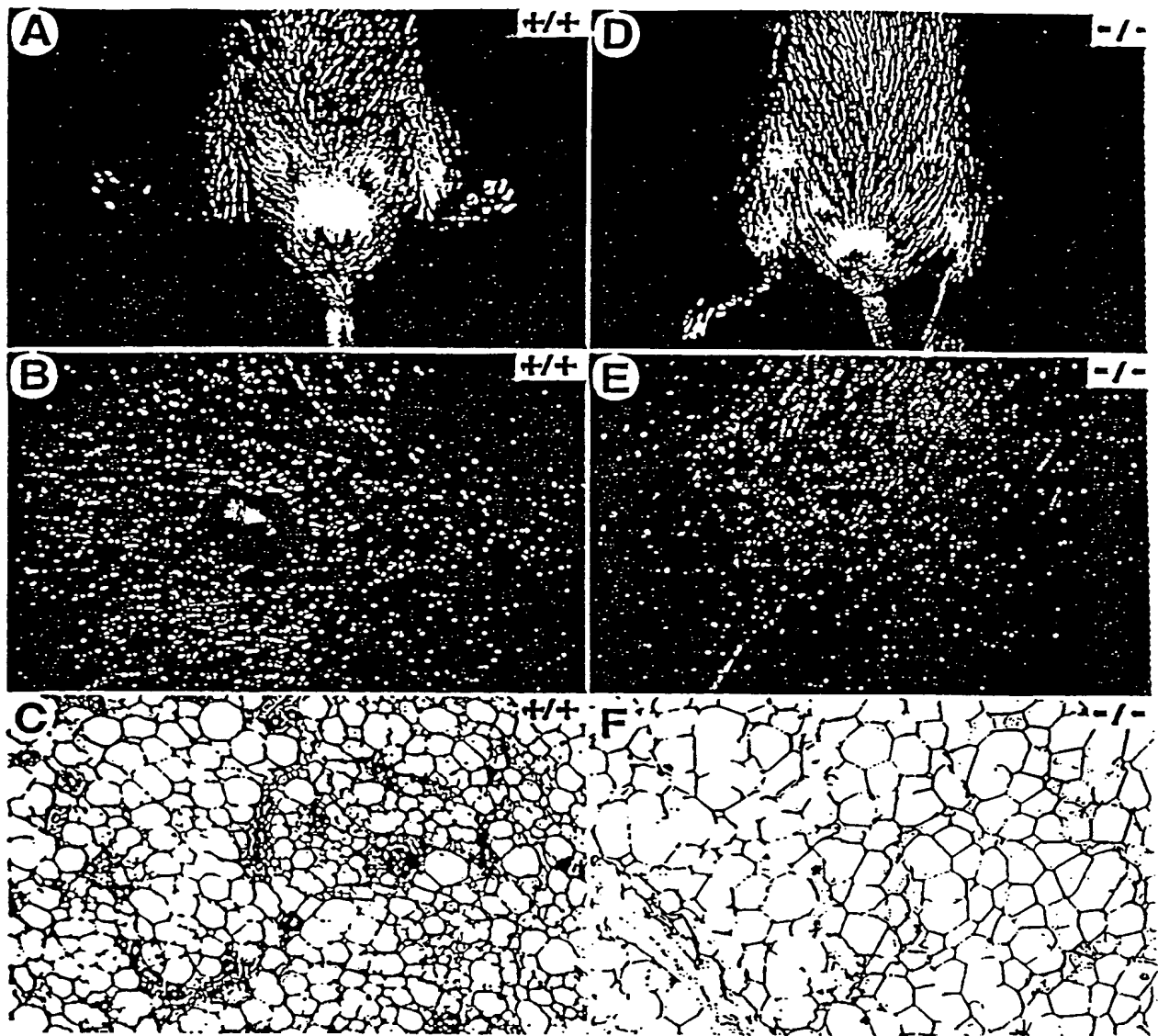
FIGS. 7A-B

9/25



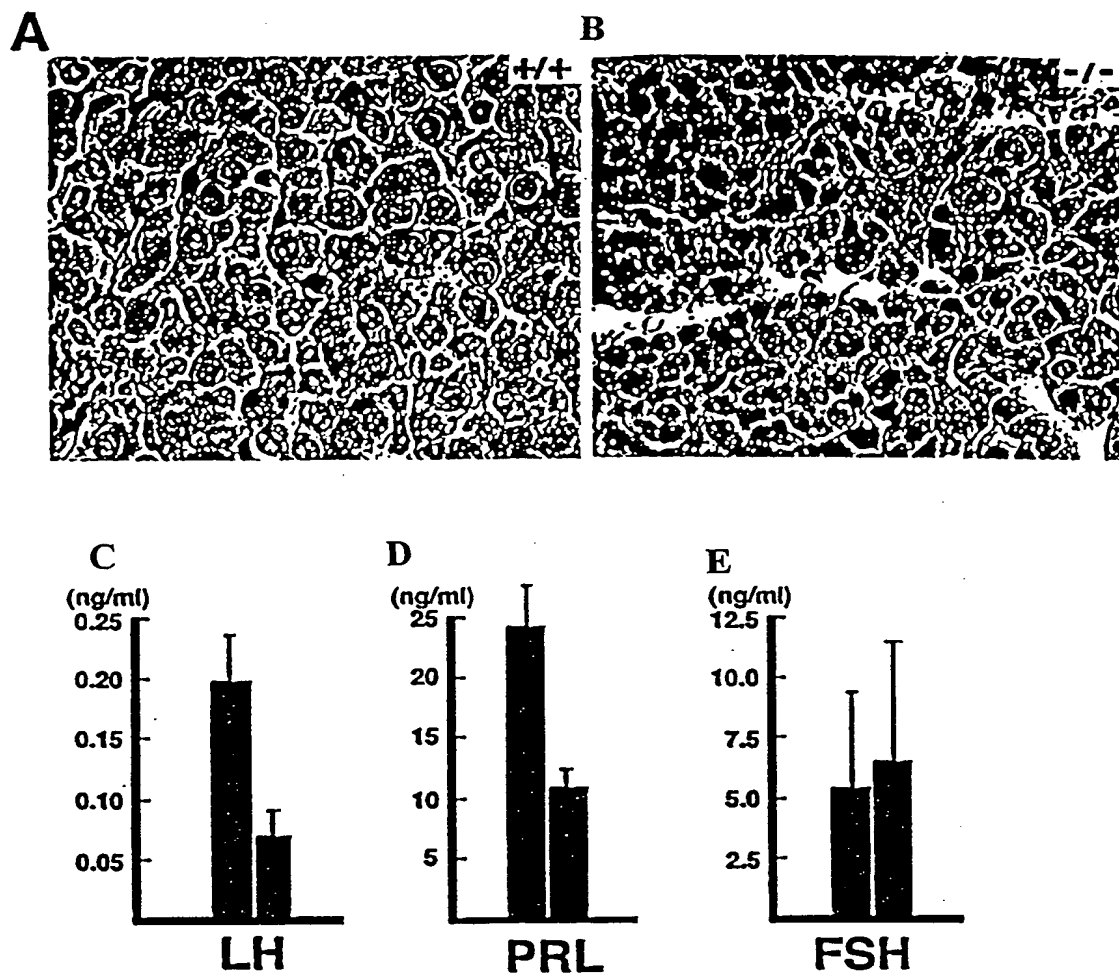
FIGS. 8A-D

10/25

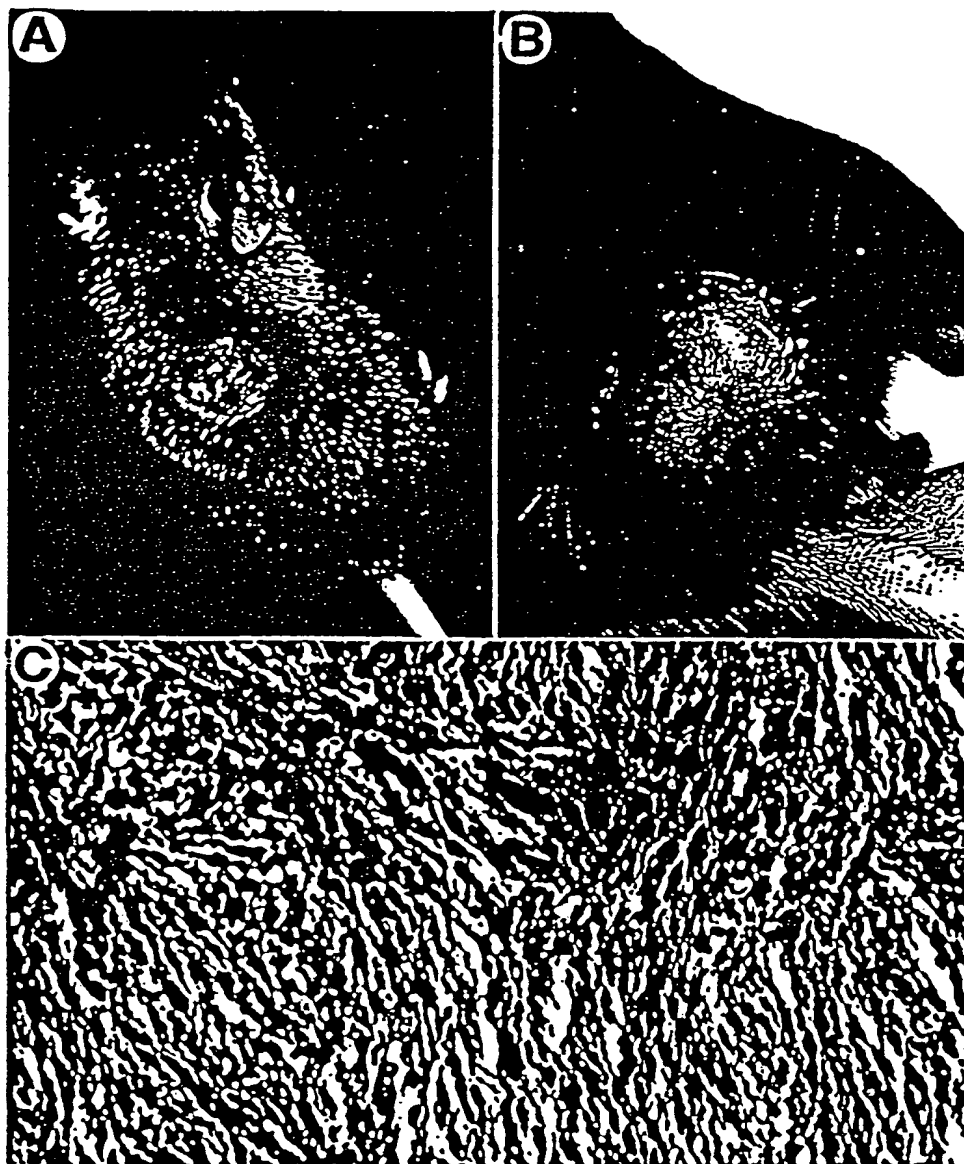


FIGS. 9A-F

11/25



FIGS. 10A-E



**FIGS. 11A-C**

13/25

10	20	30	40	50	60	70	80
ACCTTTGGGT	TGCTGGGACG	GACTCTGGCC	GCCTCAGCGT	CCGCCCTCAG	GCCCGTGGCC	GCTGTCCAGG	AGCTCTGCTC
90	100	110	120	130	140	150	160
TCCCCTCCAG	AGTTAATTAT	TTATATTGTA	AAGAATTTTA	ACAGTCCTGG	GGACTTCCTT	GAAGGATCAT	TTTCACTTTT
170	180	190	200	210	220	230	240
GCTCAGAAGA	AAGCTCTGGA	TCTATCAAAT	AAAGAAGTCC	TTCGTGTGGG	CTACATATAT	AGATGTTTTT	ATGAAGAGGA
							M K R
250	260	270	280	290	300	310	320
GTGAAAAGCC	AGAAGGATAT	AGACAAATGA	GGCCTAAGAC	CTTTCCTGCC	AGTAACTATA	CTGTCACTAG	CCGGCAAATG
S E K P	E G Y	R Q M	R P K T	F P A	S N Y	T V S S	R Q M
330	340	350	360	370	380	390	400
TTACAAGAAA	TTCGGGAATC	CCTTAGGAAT	TTATCTAAAC	CATCTGATGC	TGCTAAGGCT	GAGCATAACA	TGAGTAAAAT
L Q E	I R E S	L R N	L S K	P S D A	A K A	E H N	M S K M
410	420	430	440	450	460	470	480
GTCAACCGAA	GATCCTCGAC	AAGTCAGAAA	TCCACCCAAA	TTTGGGACGC	ATCATAAAGC	CTTGCAGGAA	ATTCGAAACT
S T E	D P R	Q V R N	P P K	F G T	H H K A	L Q E	I R N
490	500	510	520	530	540	550	560
CTCTGCTTCC	ATTTGCAAAT	GAAACAAATT	CTTCTCGGAG	TACTTCAGAA	GTTAATCCAC	AAATGCTTCA	AGACTTGCAA
S L L P	F A N	E T N	S S R S	T S E	V N P	Q M L Q	D L Q
570	580	590	600	610	620	630	640
GCTGCTGGAT	TTGATGAGGA	TATGGTTATA	CAAGCTCTTC	AGAAAACTAA	CAACAGAAGT	ATAGAAGCAG	CAATTGAATT
A A G	F D E D	M V I	Q A L	Q K T N	N R S	I E A	A I E F
650	660	670	680	690	700	710	720
CATTAGTAAA	ATGAGTTACC	AAGATCCTCG	ACGAGAGCAG	ATGGCTGCAG	CAGCTGCCAG	ACCTATTAAAT	GCCAGCATGA
I S K	M S Y	Q D P R	R E Q	M A A	A A A R	P I N	A S M
730	740	750	760	770	780	790	800
AACCAGGGAA	TGTGCAGCAA	TCAGTTAACC	GCAAACAGAG	CTGGAAGGT	TCTAAAGAAT	CCTTAGTTCC	TCAGAGGCAT
K P G N	V Q Q	S V N	R K Q S	W K G	S K E	S L V P	Q R H
810	820	830	840	850	860	870	880
GGCCCCGCCAC	TAGGAGAAAG	TGTGGCCTAT	CATTCTGAGA	GTCCCAACTC	ACAGACAGAT	GTAGGAAGAC	CTTTGTCTGG
G P P	L G E S	V A Y	H S E	S P N S	Q T D	V G R	P L S G
890	900	910	920	930	940	950	960
ATCTGGTATA	TCAGCATTTG	TTCAAGCTCA	CCCTAGCAAC	GGACAGAGAG	TGAACCCCCC	ACCACCACCT	CAAGTAAGGA
S G I	S A F	V Q A H	P S N	G Q R	V N P P	P P P	Q V R
970	980	990	1000	1010	1020	1030	1040
GTGTTACTCC	TCCACCACCT	CCAAGAGGCC	AGACTCCCCC	TCCAAGAGGT	ACAACTCCAC	CTCCCCCTTC	ATGGGAACCA
S V T P	P P P	P R G	Q T P P	P R G	T T P	P P P S	W E P

FIG. 12

1050 1060 1070 1080 1090 1100 1110 1120  
\* \* \* \* \*  
AACTCTCAAA CAAAGCGCTA TTCTGGAAAC ATGGAATACG TAATCTCCCG AATCTCTCCT GTCCACCTG GGGCATGGCA  
N S Q T K R Y S G N M E Y V I S R I S P V P P G A W Q  
1130 1140 1150 1160 1170 1180 1190 1200  
\* \* \* \* \*  
AGAGGGCTAT CCTCCACCAC CTCTCAACAC TTCCCCCATG AATCCTCCTA ATCAAGGACA GAGAGGCATT AGTTCTGTTC  
E G Y P P P P L N T S P M N P P N Q G Q R G I S S V  
1210 1220 1230 1240 1250 1260 1270 1280  
\* \* \* \* \*  
CTGTTGGCAG ACAACCAATC ATCATGCAGA GTTCTAGCAA ATTTAACTTT CCATCAGGGA GACCTGGAAT GCAGAATGGT  
P V G R Q P I I M Q S S S K F N F P S G R P G M Q N G  
1290 1300 1310 1320 1330 1340 1350 1360  
\* \* \* \* \*  
ACTGGACAAA CTGATTTTCAT GATACACCAA AATGTTGTCC CTGCTGGCAC TGTGAATCGG CAGCCACCAC CTCCATATCC  
T G Q T D F M I H Q N V V P A G T V N R Q P P P P Y P  
1370 1380 1390 1400 1410 1420 1430 1440  
\* \* \* \* \*  
TCTGACAGCA GCTAATGGAC AAAGCCCTTC TGCTTTACAA ACAGGGGGAT CTGCTGCTCC TTCGTCATAT ACAAATGGAA  
L T A A N G Q S P S A L Q T G G S A A P S S Y T N G  
1450 1460 1470 1480 1490 1500 1510 1520  
\* \* \* \* \*  
GTATTCCTCA GTCTATGATG GTGCCAAACA GAAATAGTCA TAACATGGAA CTATATAACA TTAGTGTACC TGGACTGCAA  
S I P Q S M M V P N R N S H N M E L Y N I S V P G L Q  
1530 1540 1550 1560 1570 1580 1590 1600  
\* \* \* \* \*  
ACAAATGGC CTCAGTCATC TTCTGCTCCA GCCAGTCAT CCCCAGGCAG TGGGCATGAA ATCCCTACAT GGCAACCTAA  
T N W P Q S S S A P A Q S S P S S G H E I P T W Q P N  
1610 1620 1630 1640 1650 1660 1670 1680  
\* \* \* \* \*  
CATACCAGTG AGGTCAAATT CTTTAAATAA CCCATTAGGA AATAGAGCAA GTCACCTCTGC TAATTCTCAG CCTTCTGCTA  
I P V R S N S F N N P L G N R A S H S A N S Q P S A  
1690 1700 1710 1720 1730 1740 1750 1760  
\* \* \* \* \*  
CAACAGTCAC TGCAATTACA CCAGCTCCTA TTCAACAGCC TGTGAAAAGT ATGCGTGTAT TAAAACCAGA GCTACAGACT  
T T V T A I T P A P I Q Q P V K S M R V L K P E L Q T  
1770 1780 1790 1800 1810 1820 1830 1840  
\* \* \* \* \*  
GCTTTAGCAC CTACACACCC TTCTTGATA CCACAGCCAA TTCAAACCTGT TCAACCCAGT CCTTTTCTCTG AGGGAACCGC  
A L A P T H P S W I P Q P I Q T V Q P S P F P E G T A  
1850 1860 1870 1880 1890 1900 1910 1920  
\* \* \* \* \*  
TTCAAATGTG ACTGTGATGC CACCTGTTGC TGAAGCTCCA AACTATCAAG GACCACCACC ACCCTACCCA AAACATCTGC  
S N V T V M P P V A E A P N Y Q G P P P P Y P K H L  
1930 1940 1950 1960 1970 1980 1990 2000  
\* \* \* \* \*  
TGCACCAAAA CCCATCTGTT CCTCCATACG AGTCAATCAG TAAGCCTAGC AAAGAGGATC AGCCAAGCTT GCCCAAGGAA  
L H Q N P S V P P Y E S I S K P S K E D Q P S L P K E  
2010 2020 2030 2040 2050 2060 2070 2080  
\* \* \* \* \*  
GATGAGAGTG AAAAGAGTTA TGAAAATGTT GATAGTGGGG ATAAAGAAAA GAAACAGATT ACAACTTCAC CTATTACTGT  
D E S E K S Y E N V D S G D K E K K Q I T T S P I T V

FIG. 12 (cont.)

2090 2100 2110 2120 2130 2140 2150 2160  
\* \* \* \* \*  
TAGGAAAAAC AAGAAAGATG AAGAGCGAAG GGAATCTCGT ATTCAAAGTT ATTCTCCTCA AGCATTATAA TTCTTTATGG  
R K N K K D E E R R E S R I Q S Y S P Q A F K F F M

2170 2180 2190 2200 2210 2220 2230 2240  
\* \* \* \* \*  
AGCAACATGT AGAAAATGTA CTCAAATCTC ATCAGCAGCG TCTACATCGT AAAAAACAAT TAGAGAATGA AATGATGCGG  
E Q H V E N V L K S H Q Q R L H R K K Q L E N E M M R

2250 2260 2270 2280 2290 2300 2310 2320  
\* \* \* \* \*  
GTTGGATTAT CTCAAGATGC CCAGGATCAA ATGAGAAAGA TGCTTTGCCA AAAAGAATCT AATTACATCC GTCTTAAAG  
V G L S Q D A Q D Q M R K M L C Q K E S N Y I R L K R

2330 2340 2350 2360 2370 2380 2390 2400  
\* \* \* \* \*  
GGCTAAAATG GACAAGTCTA TGTTTGTGAA GATAAAGACA CTAGGAATAG GAGCATTGCG TGAAGTCTGT CTAGCAAGAA  
A K M D K S M F V K I K T L G I G A F G E V C L A R

2410 2420 2430 2440 2450 2460 2470 2480  
\* \* \* \* \*  
AAGTAGATAC TAAGGCTTTG TATGCAACAA AAATCTCTCG AAAGAAAGAT GTTCTTCTTC GAAATCAAGT CGCTCATGTT  
K V D T K A L Y A T K T L R K K D V L L R N Q V A H V

2490 2500 2510 2520 2530 2540 2550 2560  
\* \* \* \* \*  
AAGGCTGAGA GAGATATCCT GGCTGAAGCT GACAATGAAT GGGTAGTTCTG TCTATATTAT TCATTCCAAG ATAAGGACAA  
K A E R D I L A E A D N E W V V R L Y Y S F Q D K D N

2570 2580 2590 2600 2610 2620 2630 2640  
\* \* \* \* \*  
TTTATACTTT GTAATGGACT ACATTCCTGG GGGTGATATG ATGAGCCTAT TAATTAGAAT GGGCATCTTT CCAGAAAGTC  
L Y F V M D Y I P G G D M M S L L I R M G I F P E S

2650 2660 2670 2680 2690 2700 2710 2720  
\* \* \* \* \*  
TGGCAGGATT CTACATAGCA GAACTTACCT GTGCAGTTGA AAGTGTTTCAT AAAATGGGTT TTATTCATAG AGATATTAAA  
L A R F Y I A E L T C A V E S V H K M G F I H R D I K

2730 2740 2750 2760 2770 2780 2790 2800  
\* \* \* \* \*  
CCTGATAATA TTTTGATTGA TCGTGATGGT CATATTAAAT TGAATGACTT TGGCCTCTGC ACTGGCTTCA GATGGACACA  
P D N I L I D R D G H I K L T D F G L C T G F R W T H

2810 2820 2830 2840 2850 2860 2870 2880  
\* \* \* \* \*  
CGATTCTAAG TACTATCAGA GTGGTGACCA TCCACGGCAA GATAGCATGG ATTTCAAGTAA TGAATGGGGG GATCCCTCAA  
D S K Y Y Q S G D H P R Q D S M D F S N E W G D P S

2890 2900 2910 2920 2930 2940 2950 2960  
\* \* \* \* \*  
GCTGTGATG TGGAGACAGA CTGAAGCCAT TAGAGCGGAG AGCTGCACGC CAGCACCAGC GATGTCTAGC ACATTCCTTG  
S C R C G D R L K P L E R R A A R Q H Q R C L A H S L

2970 2980 2990 3000 3010 3020 3030 3040  
\* \* \* \* \*  
GTTGGGACTC CCAATTATAT TGCACCTGAA GTGTTGCTAC GAACAGGATA CACACAGTTG TGTGATTGGT GGAGTGTGG  
V G T P N Y I A P E V L L R T G Y T Q L C D W W S V G

3050 3060 3070 3080 3090 3100 3110 3120  
\* \* \* \* \*  
TGTTATTCTT TTTGAAATGT TGGTGGGACA ACCTCCTTTC TTGGCACAAA CACCATTAGA AACACAAATG AAGGTTATCA  
V I L F E M L V G Q P P F L A Q T P L E T Q M K V I

FIG. 12 (cont.)



3130 3140 3150 3160 3170 3180 3190 3200  
\* \* \* \* \*  
ACTGGCAAAC ATCTCTTCAC ATTCCACCAC AAGCTAAACT CAGTCCTGAA GCTTCTGATC TTATTATTAA ACTTTGCCGA  
N W Q T S L H I P P Q A K L S P E A S D L I I K L C R

3210 3220 3230 3240 3250 3260 3270 3280  
\* \* \* \* \*  
GGACCCGAAG ATCGCTTAGG CAAGAATGGT GCTGATGAAA TAAAAGCTCA TCCATTTTTT AAAACAATTG ACTTCTCCAG  
G P E D R L G K N G A D E I K A H P F F K T I D F S S

3290 3300 3310 3320 3330 3340 3350 3360  
\* \* \* \* \*  
TGACCTGAGA CAGCAGTCTG CTTCATACAT TCCTAAAATC ACACACCCAA CAGATACATC AAATTTTGAT CCTGTTGATC  
D L R Q Q S A S Y I P K I T H P T D T S N F D P V D

3370 3380 3390 3400 3410 3420 3430 3440  
\* \* \* \* \*  
CTGATAAATT ATGGAGTGAT GATAACGAGG AAGAAAATGT AAATGACACT CTCAATGGAT GGTATAAAAA TGGAAAGCAT  
P D K L W S D D N E E E N V N D T L N G W Y K N G K H

3450 3460 3470 3480 3490 3500 3510 3520  
\* \* \* \* \*  
CCTGAACATG CATTCATGA ATTTACCTTC CGAAGGTTTT TTGATGACAA TGGCTACCCA TATAATTATC CGAAGCCTAT  
P E H A F Y E F T F R R F F D D N G Y P Y N Y P K P I

3530 3540 3550 3560 3570 3580 3590 3600  
\* \* \* \* \*  
TGAATATGAA TACATTAATT CACAAGGCTC AGAGCAGCAG TCGGATGAAG ATGATCAAAA CACAGGCTCA GAGATTAAAA  
E Y E Y I N S Q G S E Q Q S D E D D Q N T G S E I K

3610 3620 3630 3640 3650 3660 3670 3680  
\* \* \* \* \*  
ATCGCGATCT AGTATATGTT TAACACACTA GTAAATAAAT GTAATGAGGA TTTGTAAAAG GGCCTGAAAT GCGAGGTGTT  
N R D L V Y V

3690 3700 3710 3720 3730 3740 3750 3760  
\* \* \* \* \*  
TTGAGGTTCT GAGAGTAAAA TTATGCAAAT ATGACAGAGC TATATATGTG TGCTCTGTGT ACAATATTTT ATTTTCCTAA

3770 3780 3790 3800 3810 3820 3830 3840  
\* \* \* \* \*  
ATTATGGGAA ATCCTTTTAA AATGTTAATT TATTCCAGCC GTTTAAATCA GTATTTAGAA AAAAATTGTT ATAAGGAAAG

3850 3860 3870 3880 3890 3900 3910 3920  
\* \* \* \* \*  
TAAATTATGA ACTGAATATT ATAGTCAGTT CTTGGTACTT AAAGTACTTA AAATAAGTAG TGCTTTGTTT AAAAGGAGAA

3930 3940 3950 3960 3970 3980  
\* \* \* \* \*  
ACCTGGTATC TATTTGTATA TATGCTAAAT AATTTTAAAA TACAAGAGTT TTTGAAATTT TTTT

FIG. 12 (cont.)

10 20 30 40 50 60 70 80  
\* \* \* \* \*  
GTGCAACATT CAATTAACCG AAAACAAAGC TGGAAAGGTT CTAAGAGATC TCTAGTTCCT CAGAGACACG GCCCATCTCT  
V Q H S I N R K Q S W K G S K E S L V P Q R H G P S L  
90 100 110 120 130 140 150 160  
\* \* \* \* \*  
AGGAGAAAAT GTGGTTTATC GTTCTGAAAG CCCCAACTCA CAGGCGGATG TAGGAAGACC TCTGTCTGGA TCCGGCATTG  
G E N V V Y R S E S P N S Q A D V G R P L S G S G I  
170 180 190 200 210 220 230 240  
\* \* \* \* \*  
CAGCATTTCG TCAAGCTCAC CCAAGCAATG GACAGAGAGT GAACCCCCCA CCACCACCTC AAGTTAGGAG TGTACTCTCT  
A A F A Q A H P S N G Q R V N P P P P P Q V R S V T P  
250 260 270 280 290 300 310 320  
\* \* \* \* \*  
CCACCACCTC CGAGAGGCCA GACCCACCT CCCCAGGCA CCACTCCCCC TCCCCCTCA TGGGAACCAA GCTCTCAGAC  
P P P P R G Q T P P P R G T T P P P P S W E P S S Q T  
330 340 350 360 370 380 390 400  
\* \* \* \* \*  
AAAGCGCTAC TCTGGGAACA TGGAGTACGT AATCTCCCGA ATCTCCCCTG TTCCACCTGG GCGGTGGCAG GAGGGGTACC  
K R Y S G N M E Y V I S R I S P V P P G A W Q E G Y  
410 420 430 440 450 460 470 480  
\* \* \* \* \*  
CTCCACCACC TCTTACCCT TCTCCCATGA ATCCCCCTAG CCAGGCTCAG AGGGCCATTA GTTCTGTTCC AGTTGGTAGA  
P P P P L T T S P M N P P S Q A Q R A I S S V P V G R  
490 500 510 520 530 540 550 560  
\* \* \* \* \*  
CAACCCATCA TCATGCAGAG TACTAGCAAA TTAACTTTA CACCAGGGCG ACCTGGAGTT CAGAATGGTG GTGGTCAGTC  
Q P I I M Q S T S K F N F T P G R P G V Q N G G G Q S  
570 580 590 600 610 620 630 640  
\* \* \* \* \*  
TGATTTTATC GTGCACCAA ATGTCCCCAC TGGTTCTGTG ACTCGGCAGC CACCACCTCC ATATCTCTG ACCCCAGCTA  
D F I V H Q N V P T G S V T R Q P P P P Y P L T P A  
650 660 670 680 690 700 710 720  
\* \* \* \* \*  
ATGGACAAAG CCCCTCTGCT TTACAAACAG GGGCTTCTGC TGCTCCACCA TCATTGCGCA ATGGAAACGT TCCTCAGTCG  
N G Q S P S A L Q T G A S A A P P S F A N G N V P Q S  
730 740 750 760 770 780 790 800  
\* \* \* \* \*  
ATGATGGTGC CCAACAGGAA CAGTCATAAC ATGGAGCTTT ATAATATTAA TGTCCTGGA CTGCAAACAG CCTGGCCCCA  
M M V P N R N S H N M E L Y N I N V P G L Q T A W P Q  
810 820 830 840 850 860 870 880  
\* \* \* \* \*  
GTCGTCTTCT GCTCCTGCGC AGTCATCCCC AAGCGGTGGG CATGAAATTC CTACATGGCA ACCTAACATA CCAGTGAGGT  
S S S A P A Q S S P S G G H E I P T W Q P N I P V R  
890 900 910 920 930 940 950 960  
\* \* \* \* \*  
CAAATTCTTT TAATAACCCA TTAGGAAGTA GAGCAAGTCA CTCTGCTAAT TCTCAGCCTT CTGCCACTAC AGTCACTGCC  
S N S F N N P L G S R A S H S A N S Q P S A T T V T A

FIG. 13

970	980	990	1000	1010	1020	1030	1040
ATCACACCCG	CTCCTATTCA	ACAGCCCGTG	AAAAGCATGC	GCGTCCTGAA	ACCAGAGCTG	CAGACTGCTT	TAGCCCCAAC
I T P	A P I Q	Q P V	K S M	R V L K	P E L	Q T A	L A P T
1050	1060	1070	1080	1090	1100	1110	1120
CCATCCTTCT	TGGATGCCAC	AGCCAGTTCA	GACTGTTTCAG	CCTACCCCTT	TTTCTGAGGG	TACAGCTTCA	AGTGTGCCTG
H P S	W M P	Q P V Q	T V Q	P T P	F S E G	T A S	S V P
1130	1140	1150	1160	1170	1180	1190	1200
TCATCCCAACC	TGTTGCTGAA	GCTCCAAGCT	ATCAAGGTCC	ACCACCGCCT	TATCCAAAAC	ATCTGCTACA	CCAAAACCCA
V I P P	V A E	A P S	Y Q G P	P P P	Y P K	H L L H	Q N P
1210	1220	1230	1240	1250	1260	1270	1280
TCTGTCCCTC	CATATGAGTC	AGTAAGTAAG	CCCTGCAAAG	ATGAACAGCC	TAGCTTACCC	AAGGAAGATG	ATAGTGAGAA
S V P	P Y E S	V S K	P C K	D E Q P	S L P	K E D	D S E K
1290	1300	1310	1320	1330	1340	1350	1360
GAGTGCGGAC	AGTGGTGACT	CTGGGGATAA	AGAAAAGAAA	CAGATTACAA	CTTCACCTAT	CACTGTTTCGG	AAAAACAAGA
S A D	S G D	S G D K	E K K	Q I T	T S P I	T V R	K N K
1370	1380	1390	1400	1410	1420	1430	1440
AAGATGAAGA	ACGAAGAGAG	TCTCGGATTC	AGAGTTACTC	CCCACAGGCC	TTTAAGTTCT	TCATGGAGCA	GCACGTAGAG
K D E E	R R E	S R I	Q S Y S	P Q A	F K F	F M E Q	H V E
1450	1460	1470	1480	1490	1500	1510	1520
AACGTCCTGA	AGTCTCATCA	GCAGCGTCTG	CATCGGAAGA	AGCAGCTAGA	AAATGAAATG	ATGCGGGTTG	GATTATCTCA
N V L	K S H Q	Q R L	H R K	K Q L E	N E M	M R V	G L S Q
1530	1540	1550	1560	1570	1580	1590	1600
AGATGCCCAG	GATCAAATGA	GAAAGATGCT	TTGCCAGAAA	GAGTCTAACT	ATATTGCTCT	TAAAAGGGCT	AAAATGGACA
D A Q	D Q M	R K M L	C Q K	E S N	Y I R L	K R A	K M D
1610	1620	1630	1640	1650	1660	1670	1680
AGTCTATGTT	TGTAAAGATA	AAGACATTAG	GAATAGGAGC	GTTTGGTGAA	GTCTGTCTAG	CAAGAAAAGT	CGATACTAAA
K S M F	V K I	K T L	G I G A	F G E	V C L	A R K V	D T K
1690	1700	1710	1720	1730	1740	1750	1760
GCTTGTATG	CAACAAAGAC	TCTTCGAAAG	AAAGACGTTT	TGCTCCGAAA	TCAGGTGGCT	CATGTGAAAG	CGGAGAGGGA
A L Y	A T K T	L R K	K D V	L L R N	Q V A	H V K	A E R D
1770	1780	1790	1800	1810	1820	1830	1840
TATCCTAGCA	GAAGCCGACA	ATGAGTGGGT	GGTCCGCCTG	TACTACTCTT	TCCAGGACAA	GGACAACCTG	TACTTTGTGA
I L A	E A D	N E W V	V R L	Y Y S	F Q D K	D N L	Y F V
1850	1860	1870	1880	1890	1900	1910	1920
TGGACTACAT	TCCTGGGGGG	GATATGATGA	GCCTATTAAT	TAGAATGGGC	ATCTTTCCTG	AAAATCTGGC	ACGATTCTAC
M D Y I	P G G	D M M	S L L I	R M G	I F P	E N L A	R F Y
1930	1940	1950	1960	1970	1980	1990	2000

FIG. 13 (cont.)

ATAGCAGAAC	TTACCTGTGC	AGTTGAAAGT	GTTTCATAAAA	TGGGTTTTTAT	TCATAGAGAT	ATTAAACCTG	ATAACATTTT
I A E	L T C A	V E S	V H K	M G F I	H R D	I K P	D N I L
2010	2020	2030	2040	2050	2060	2070	2080
*	*	*	*	*	*	*	*
GATTGACCGT	GATGGCCATA	TTAAATTGAC	TGACTTTGGC	TTGTGCACTG	GCTTCAGATG	GACACATGAC	TCCAAGTACT
I D R	D G H	I K L T	D F G	L C T	G F R W	T H D	S K Y
2090	2100	2110	2120	2130	2140	2150	2160
*	*	*	*	*	*	*	*
ACCAGAGTGG	GGATCACCCA	CGGCAAGATA	GCATGGATTT	CAGTAACGAA	TGGGGAGATC	CTTCCAATTG	TCGGTGTGGG
Y Q S G	D H P	R Q D	S M D F	S N E	W G D	P S N C	R C G
2170	2180	2190	2200	2210	2220	2230	2240
*	*	*	*	*	*	*	*
GACAGACTGA	AGCCACTGGA	GCGGAGAGCT	GCTCGCCAGC	ACCAGCGATG	TCTAGCCCAT	TCTCTGGTTG	GGACTCCCAA
D R L	K P L E	R R A	A R Q	H Q R C	L A H	S L V	G T P N
2250	2260	2270	2280	2290	2300	2310	2320
*	*	*	*	*	*	*	*
TTATATTGCA	CCTGAAGTGC	TACTGCGAAC	AGGATATACA	CAGCTGTGTG	ACTGGTGGAG	TGTTGGTGTG	ATTCTTTGTG
Y I A	P E V	L L R T	G Y T	Q L C	D W W S	V G V	I L C
2330	2340	2350	2360	2370	2380	2390	2400
*	*	*	*	*	*	*	*
AAATGTTGGT	GGGACAACCT	CCTTTCTTGG	CACAAACCCC	ATTAGAAACA	CAAATGAAGG	TTATCATCTG	GCAAACCTCT
E M L V	G Q P	P F L	A Q T P	L E T	Q M K	V I I W	Q T S
2410	2420	2430	2440	2450	2460	2470	2480
*	*	*	*	*	*	*	*
CTACACATCC	CTCCTCAAGC	TAAGCTGAGT	CCTGAAGCCT	CTGACCTCAT	TATCAAAGTG	TGTCGAGGAC	CAGAAGACCG
L H I	P P Q A	K L S	P E A	S D L I	I K L	C R G	P E D R
2490	2500	2510	2520	2530	2540	2550	2560
*	*	*	*	*	*	*	*
CCTCGGCAAG	AACGGTGTCTG	ATGAGATAAA	GGCTCATCCA	TTTTTTAAGA	CCATCGATTT	CTCTAGTGAT	CTGAGACAGC
L G K	N G A	D E I K	A H P	F F K	T I D F	S S D	L R Q
2570	2580	2590	2600	2610	2620	2630	2640
*	*	*	*	*	*	*	*
AGTCTGCTTC	ATACATCCCT	AAAATCACGC	ATCCAACAGA	TACATCCAAT	TTCGACCCTG	TTGATCCTGA	TAAATGTGG
Q S A S	Y I P	K I T	H P T D	T S N	F D P	V D P D	K L W
2650	2660	2670	2680	2690	2700	2710	2720
*	*	*	*	*	*	*	*
AGCGATGGCA	GCGAGGAGGA	AAATATCAGT	GACACTCTGA	GCGGATGGTA	TAAAAATGGG	AAGCACCCCG	AGCACGCTTT
S D G	S E E E	N I S	D T L	S G W Y	K N G	K H P	E H A F
2730	2740	2750	2760	2770	2780	2790	2800
*	*	*	*	*	*	*	*
CTATGAGTTC	ACCTTTTCGGA	GTTTTTTTGA	TGACAATGGC	TACCCATATA	ATTATCCAAA	GCCTATTGAG	TATGAATACA
Y E F	T F R	R F F D	D N G	Y P Y	N Y P K	P I E	Y E Y
2810	2820	2830	2840	2850	2860	2870	2880
*	*	*	*	*	*	*	*
TTCATTACAA	GGGCTCAGAA	CAACAGTCTG	ATGAAGATGA	TCAACACACA	AGCTCCGATG	GAAACAACCG	AGATCTAGTG
I H S Q	G S E	Q Q S	D E D D	Q H T	S S D	G N N R	D L V
2890	2900	2910	2920	2930	2940	2950	2960
*	*	*	*	*	*	*	*
TATGTTTAAT	AAACTAGGAG	ATCATTGTAA	GAATTTGCAA	GAGGCCTGAA	GTGCAGGGGT	TTTTGAAGTT	TTGAGAAAA
Y V	*	*	*	*	*	*	*

FIG. 13 (cont.)

2970	2980	2990	3000	3010	3020	3030	3040
*	*	*	*	*	*	*	*
TATGCAAATG	TGACAGAGTT	TGTGTGCTCT	GTGTACAATA	TTTTATTTTC	CTAAGTTATG	GGAAATTGTT	TTAAAATGTT
3050	3060	3070	3080	3090	3100	3110	3120
*	*	*	*	*	*	*	*
AATTTATTCC	ACCCTTTTAA	TTCAAGTAATT	TAGAAAAAAT	TGTTATAAGG	AAAGTAAATT	ATGAACTGAG	TATTATAGTC
3130	3140	3150	3160	3170	3180	3190	3200
*	*	*	*	*	*	*	*
AATTCTTGGT	ACTTAAAGTA	CTTAAAAAGA	GAAGCCTGGT	ATCTTTTGTA	TATATAATAA	ATAATTTTAA	AATCCCAAAA
3210							
*							
AAAAAAAAAA	AAA						

FIG. 13 (cont.)

10 20 30 40 50 60 70 80  
\* \* \* \* \*  
ATGAGAGCCA CCCCGAAGTT TGGACCTTAT CAAAAAGCTC TCAGGGAAT CCGATATTCC CTCCTGCCTT TTGCCAACGA  
M R A T P K F G P Y Q K A L R E I R Y S L L P F A N E  
90 100 110 120 130 140 150 160  
\* \* \* \* \*  
GTCAGGCACT TCGGCAGCTG CAGAGGTGAA CCGGCAGATG CTTCAGGAGT TGGTGAATGC GGCATGTGAC CAGGAGATGG  
S G T S A A A E V N R Q M L Q E L V N A A C D Q E M  
170 180 190 200 210 220 230 240  
\* \* \* \* \*  
CTGGCAGAGC GCTCACGCAG ACGGGCAGTA GGAGTATCGA AGCTGCCTTG GAGTACATCA GTAAGATGGG CTACCTGGAC  
A G R A L T Q T G S R S I E A A L E Y I S K M G Y L D  
250 260 270 280 290 300 310 320  
\* \* \* \* \*  
CCCAGGAATG AGCAGATTGT GCGAGTCATC AAGCAGACCT CCCCAGGAAA GGGCCTGGCG TCCACCCCGG TGA CTGCGG  
P R N E Q I V R V I K Q T S P G K G L A S T P V T R R  
330 340 350 360 370 380 390 400  
\* \* \* \* \*  
GCCCAGTTTC GAGGGCACAG GGGAAAGCACT CCCATCCTAC CACCAGCTGG GTGGTGCAAA CTACGAGGGC CCCGCCGCAC  
P S F E G T G E A L P S Y H Q L G G A N Y E G P A A  
410 420 430 440 450 460 470 480  
\* \* \* \* \*  
TGGAGGAGAT GCCGCGGCAA TATTTAGACT TTCTCTTCCC TGGAGCCGGA GCCGGCACCC ACGGTGCCCA GGCTCACCAG  
L E E M P R Q Y L D F L F P G A G A G T H G A Q A H Q  
490 500 510 520 530 540 550 560  
\* \* \* \* \*  
CATCCTCCCA AAGGGTACAG CACAGCAGTA GAGCCAAGTG CGCACTTTCC GGGCACACAC TATGGTCTGT GTCATCTACT  
H P P K G Y S T A V E P S A H F P G T H Y G R G H L L  
570 580 590 600 610 620 630 640  
\* \* \* \* \*  
ATCGGAGCAG TCTGGGTATG GGGTGCAGCG CAGTTCTCTCC TTCCAGAACA AGACGCCACC AGATGCCTAT TCCAGCATGG  
S E Q S G Y G V Q R S S S F Q N K T P P D A Y S S M  
650 660 670 680 690 700 710 720  
\* \* \* \* \*  
CCAAGGCCCA GGGTGGCCCT CCGCCAGCC TCACCTTTCC TGCCCATGCT GGGCTGTACA CTGCCTCGCA CCACAAGCCG  
A K A Q G G P P A S L T F P A H A G L Y T A S H H K P  
730 740 750 760 770 780 790 800  
\* \* \* \* \*  
GCGGCTACCC CACCTGGGGC CCACCCATTA CATGTGTTGG GCACCCGGGG TCCCACGTTT ACTGGCGAAA GCTCTGCACA  
A A T P P G A H P L H V L G T R G P T F T G E S S A Q  
810 820 830 840 850 860 870 880  
\* \* \* \* \*  
GGCTGTGCTG GCACCGTCCA GGAACAGCCT CAATGCTGAC TTGTACGAGC TGGGCTCCAC GGTGCCCTGG TCTGCAGCTC  
A V L A P S R N S L N A D L Y E L G S T V P W S A A  
890 900 910 920 930 940 950 960  
\* \* \* \* \*  
CACTGGCACG CCGCGACTCG CTGCAGAAGC AGGGTCTAGA AGCCTCGCGG CCGCATGTGG CTTTTCGGGC TGGCCCCAGC  
P L A R R D S L Q K Q G L E A S R P H V A F R A G P S

FIG. 14

22/25

970	980	990	1000	1010	1020	1030	1040
AGGACCAACT	CCTTCAACAA	CCCACAACCT	GAGCCCTCAC	TGCCCGCCCC	CAACACGGTC	ACCGCCGTGA	CGGCCGCACA
R T N	S F N N	P Q P	E P S	L P A P	N T V	T A V	T A A H
1050	1060	1070	1080	1090	1100	1110	1120
CATCCTTCAC	CCTGTGAAGA	GCGTGGGTGT	GCTGCGGCCC	GAGCCCCAGA	CAGCCGTGGG	GCCCTCGCAC	CCCCCTGGG
I L H	P V K	S V R V	L R P	E P Q	T A V G	P S H	P A W
1130	1140	1150	1160	1170	1180	1190	1200
TGGCTGCGCC	CACAGCACCT	GCCACTGAGA	GCCTGGAGAC	GAAGGAGGGC	AGCGCAGGCC	CACACCCGCT	GGATGTGGAC
V A A P	T A P	A T E	S L E T	K E G	S A G	P H P L	D V D
1210	1220	1230	1240	1250	1260	1270	1280
TATGGCGGCT	CCGAGCGCAG	GTGCCCACCG	CCTCCGTATC	CAAAGCACTT	GCTGCTGCCC	AGTAAGTCTG	AGCAGTACAG
Y G G	S E R R	C P P	P P Y	P K H L	L L P	S K S	E Q Y S
1290	1300	1310	1320	1330	1340	1350	1360
CGTGGACCTG	GACAGCCTGT	GCACCAGTGT	GCAGCAGAGT	CTGCGAGGGG	GCACTGATCT	AGACGGGAGT	GACAAGAGCC
V D L	D S L	C T S V	Q Q S	L R G	G T D L	D G S	D K S
1370	1380	1390	1400	1410	1420	1430	1440
ACAAAGGTGC	GAAGGGAGAC	AAAGCTGGCA	GAGACAAAAA	GCAGATTTCAG	ACCTCCCCGG	TGCCTGTCCG	CAAGAATAGC
H K G A	K G D	K A G	R D K K	Q I Q	T S P	V P V R	K N S
1450	1460	1470	1480	1490	1500	1510	1520
AGAGATGAAG	AGAAGAGAGA	GTCTCGCATC	AAGAGTTACT	CCCCTTATGC	CTTCAAATTC	TTCATGGAGC	AACACGTGGA
R D E	E K R E	S R I	K S Y	S P Y A	F K F	F M E	Q H V E
1530	1540	1550	1560	1570	1580	1590	1600
GAATGTCATC	AAAACCTACC	AGCAGAAGGT	CAGCCGGAGG	CTACAGCTGG	AGCAGGAAAT	GGCCAAAGCT	GGGCTCTGTG
N V I	K T Y	Q Q K V	S R R	L Q L	E Q E M	A K A	G L C
1610	1620	1630	1640	1650	1660	1670	1680
AGGCCGAGCA	GGAGCAGATG	AGGAAGATCC	TCTACCAGAA	GGAGTCTAAC	TACAACCGGC	TGAAGAGGGC	CAAGATGGAC
E A E Q	E Q M	R K I	L Y Q K	E S N	Y N R	L K R A	K M D
1690	1700	1710	1720	1730	1740	1750	1760
AAGTCCATGT	TTGTGAAAAT	CAAGACTCTA	GGCATCGGTG	CCTTTGGGGA	AGTGTGCCTC	GCTTGTAAGC	TGGACACTCA
K S M	F V K I	K T L	G I G	A F G E	V C L	A C K	L D T H
1770	1780	1790	1800	1810	1820	1830	1840
CGCTCTGTAC	GCCATGAAGA	CTCTCAGGAA	GAAGGATGTC	CTGAACCGGA	ATCAAGTGGC	CCATGTCAAG	GCTGAGAGGG
A L Y	A M K	T L R K	K D V	L N R	N Q V A	H V K	A E R
1850	1860	1870	1880	1890	1900	1910	1920
ACATCCTGGC	TGAAGCAGAC	AATGAGTGGG	TGGTCAAACCT	CTACTACTCC	TTCCAGGACA	AGGACAGCCT	GTACTTTGTG
D I L A	E A D	N E W	V V K L	Y Y S	F Q D	K D S L	Y F V
1930	1940	1950	1960	1970	1980	1990	2000

FIG. 14 (cont.)

970 980 990 1000 1010 1020 1030 1040  
\* \* \* \* \*  
AGGACCAACT CCTTCAACAA CCCACAACCT GAGCCCTCAC TGCCCGCCCC CAACACGGTC ACCGCCGTGA CGGCCGCACA  
R T N S F N N P Q P E P S L P A P N T V T A V T A A H

1050 1060 1070 1080 1090 1100 1110 1120  
\* \* \* \* \*  
CATCCTTCAC CCTGTGAAGA GCGTGCGTGT GCTGCGGCCC GAGCCCCAGA CAGCCGTGGG GCCCTGCGAC CCCGCCCTGGG  
I L H P V K S V R V L R P E P Q T A V G P S H P A W

1130 1140 1150 1160 1170 1180 1190 1200  
\* \* \* \* \*  
TGGCTGCGCC CACAGCACCT GCCACTGAGA GCCTGGAGAC GAAGGAGGGC AGCGCAGGCC CACACCCGCT GGATGTGGAC  
V A A P T A P A T E S L E T K E G S A G P H P L D V D

1210 1220 1230 1240 1250 1260 1270 1280  
\* \* \* \* \*  
TATGGCGGCT CCGAGCGCAG GTGCCCACCG CCTCCGTATC CAAAGCACTT GCTGCTGCCC AGTAAGTCTG AGCAGTACAG  
Y G G S E R R C P P P P Y P K H L L L P S K S E Q Y S

1290 1300 1310 1320 1330 1340 1350 1360  
\* \* \* \* \*  
CGTGGACCTG GACAGCCTGT GCACCAAGTGT GCAGCAGAGT CTGCGAGGGG GCACTGATCT AGACGGGAGT GACAAGAGCC  
V D L D S L C T S V Q Q S L R G G T D L D G S D K S

1370 1380 1390 1400 1410 1420 1430 1440  
\* \* \* \* \*  
ACAAAGGTGC GAAGGGAGAC AAAGCTGGCA GAGACAAAA GCAGATTCAG ACCTCCCCGG TGCTGTCCG CAAGAATAGC  
H K G A K G D K A G R D K K Q I Q T S P V P V R K N S

1450 1460 1470 1480 1490 1500 1510 1520  
\* \* \* \* \*  
AGAGATGAAG AGAAGAGAGA GTCTCGCATC AAGAGTTACT CCCCTTATGC CTTCAAATTC TTCATGGAGC AACACGTGGA  
R D E E K R E S R I K S Y S P Y A F K F F M E Q H V E

1530 1540 1550 1560 1570 1580 1590 1600  
\* \* \* \* \*  
GAATGTCATC AAAACCTACC AGCAGAAGGT CAGCCGAGG CTACAGCTGG AGCAGGAAAT GGCCAAAGCT GGGCTCTGTG  
N V I K T Y Q Q K V S R R L Q L E Q E M A K A G L C

1610 1620 1630 1640 1650 1660 1670 1680  
\* \* \* \* \*  
AGGCCGAGCA GGAGCAGATG AGGAAGATCC TCTACCAGAA GGAGTCTAAC TACAACCGGC TGAAGAGGGC CAAGATGGAC  
E A E Q E Q M R K I L Y Q K E S N Y N R L K R A K M D

1690 1700 1710 1720 1730 1740 1750 1760  
\* \* \* \* \*  
AAGTCCATGT TTGTGAAAAT CAAGACTCTA GGCATCGGTG CCTTTGGGGA AGTGTGCCTC GCTTGTAAGC TGGACACTCA  
K S M F V K I K T L G I G A F G E V C L A C K L D T H

1770 1780 1790 1800 1810 1820 1830 1840  
\* \* \* \* \*  
CGCTCTGTAC GCCATGAAGA CTCTCAGGAA GAAGGATGTC CTGAACCGGA ATCAAGTGGC CCATGTCAAG GCTGAGAGGG  
A L Y A M K T L R K K D V L N R N Q V A H V K A E R

1850 1860 1870 1880 1890 1900 1910 1920  
\* \* \* \* \*  
ACATCCTGGC TGAAGCAGAC AATGAGTGGG TGGTCAAACCT CTACTACTCC TTCCAGGACA AGGACAGCCT GTACTTTGTG  
D I L A E A D N E W V V K L Y Y S F Q D K D S L Y F V

1930 1940 1950 1960 1970 1980 1990 2000  
\* \* \* \* \*  
\* \* \* \* \*

FIG. 14 (cont.)



2970	2980	2990	3000	3010	3020	3030	3040
CTCGAGGAAA	CCCAAAATGA	GATTTCTTTT	CAGAAGACAA	ACTCAAGCTT	AGGAATCCTT	CATTTT TAGT	TCTGGTAAAT
3050	3060	3070	3080	3090	3100	3110	3120
GGGCAACAGG	AAGAGTCAAC	ATGATT TCAA	ATTAGCCCTC	TGAGGACCTT	CACTGCATTA	AAACAGTATT	TTTTAAAAAA
3130	3140	3150					
TTAGTACAGT	ATGGAAAGAG	CACTTATTTT	GGGGG				

**FIG. 14 (cont.)**

[illegible]

FIG. 15

## SEQUENCE LISTING

&lt;110&gt; Yale University

<120> TREATMENT AND PREVENTION OF CANCER AND PITUITARY  
DISORDERS WITH LATS PROTEINS, DERIVATIVES AND  
FRAGMENTS, AND LATS KNOCK-OUT ANIMAL MODELS

&lt;130&gt; 6523-020-228

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 8

&lt;170&gt; PatentIn Ver. 2.0

&lt;210&gt; 1

&lt;211&gt; 3984

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (231) .. (3620)

&lt;400&gt; 1

```

acctttgggt tgctgggacg gactctggcc gcctcagcgt ccgccctcag gcccggtggcc 60
gctgtccagg agctctgctc tcccctccag agttaattat ttatattgta aagaatttta 120
acagtcctgg ggacttcctt gaaggatcat tttcactttt gctcagaaga aagctctgga 180
tctatcaaat aaagaagtcc ttcgtgtggg ctacatatat agatgttttc atg aag      236
                                   Met Lys
                                   1

agg agt gaa aag cca gaa gga tat aga caa atg agg cct aag acc ttt      284
Arg Ser Glu Lys Pro Glu Gly Tyr Arg Gln Met Arg Pro Lys Thr Phe
      5                      10                      15

cct gcc agt aac tat act gtc agt agc cgg caa atg tta caa gaa att      332
Pro Ala Ser Asn Tyr Thr Val Ser Ser Arg Gln Met Leu Gln Glu Ile
      20                      25                      30

cgg gaa tcc ctt agg aat tta tct aaa cca tct gat gct gct aag gct      380
Arg Glu Ser Leu Arg Asn Leu Ser Lys Pro Ser Asp Ala Ala Lys Ala
      35                      40                      45                      50

gag cat aac atg agt aaa atg tca acc gaa gat cct cga caa gtc aga      428
Glu His Asn Met Ser Lys Met Ser Thr Glu Asp Pro Arg Gln Val Arg
      55                      60                      65

aat cca ccc aaa ttt ggg acg cat cat aaa gcc ttg cag gaa att cga      476
Asn Pro Pro Lys Phe Gly Thr His His Lys Ala Leu Gln Glu Ile Arg
      70                      75                      80

```

aac tct ctg ctt cca ttt gca aat gaa aca aat tct tct cgg agt act	524
Asn Ser Leu Leu Pro Phe Ala Asn Glu Thr Asn Ser Ser Arg Ser Thr	
85 90 95	
tca gaa gtt aat cca caa atg ctt caa gac ttg caa gct gct gga ttt	572
Ser Glu Val Asn Pro Gln Met Leu Gln Asp Leu Gln Ala Ala Gly Phe	
100 105 110	
gat gag gat atg gtt ata caa gct ctt cag aaa act aac aac aga agt	620
Asp Glu Asp Met Val Ile Gln Ala Leu Gln Lys Thr Asn Asn Arg Ser	
115 120 125 130	
ata gaa gca gca att gaa ttc att agt aaa atg agt tac caa gat cct	668
Ile Glu Ala Ala Ile Glu Phe Ile Ser Lys Met Ser Tyr Gln Asp Pro	
135 140 145	
cga cga gag cag atg gct gca gca gct gcc aga cct att aat gcc agc	716
Arg Arg Glu Gln Met Ala Ala Ala Ala Arg Pro Ile Asn Ala Ser	
150 155 160	
atg aaa cca ggg aat gtg cag caa tca gtt aac cgc aaa cag agc tgg	764
Met Lys Pro Gly Asn Val Gln Gln Ser Val Asn Arg Lys Gln Ser Trp	
165 170 175	
aaa ggt tct aaa gaa tcc tta gtt cct cag agg cat ggc ccg cca cta	812
Lys Gly Ser Lys Glu Ser Leu Val Pro Gln Arg His Gly Pro Pro Leu	
180 185 190	
gga gaa agt gtg gcc tat cat tct gag agt ccc aac tca cag aca gat	860
Gly Glu Ser Val Ala Tyr His Ser Glu Ser Pro Asn Ser Gln Thr Asp	
195 200 205 210	
gta gga aga cct ttg tct gga tct ggt ata tca gca ttt gtt caa gct	908
Val Gly Arg Pro Leu Ser Gly Ser Gly Ile Ser Ala Phe Val Gln Ala	
215 220 225	
cac cct agc aac gga cag aga gtg aac ccc cca cca cca cct caa gta	956
His Pro Ser Asn Gly Gln Arg Val Asn Pro Pro Pro Pro Pro Gln Val	
230 235 240	
agg agt gtt act cct cca cca cct cca aga ggc cag act ccc cct cca	1004
Arg Ser Val Thr Pro Pro Pro Pro Pro Arg Gly Gln Thr Pro Pro Pro	
245 250 255	
aga ggt aca act cca cct ccc cct tca tgg gaa cca aac tct caa aca	1052
Arg Gly Thr Thr Pro Pro Pro Pro Ser Trp Glu Pro Asn Ser Gln Thr	
260 265 270	
aag cgc tat tct gga aac atg gaa tac gta atc tcc cga atc tct cct	1100
Lys Arg Tyr Ser Gly Asn Met Glu Tyr Val Ile Ser Arg Ile Ser Pro	
275 280 285 290	
gtc cca cct ggg gca tgg caa gag ggc tat cct cca cca cct ctc aac	1148
Val Pro Pro Gly Ala Trp Gln Glu Gly Tyr Pro Pro Pro Pro Leu Asn	
295 300 305	

act tcc ccc atg aat cct cct aat caa gga cag aga ggc att agt tct	1196
Thr Ser Pro Met Asn Pro Pro Asn Gln Gly Gln Arg Gly Ile Ser Ser	
310 315 320	
gtt cct gtt ggc aga caa cca atc atc atg cag agt tct agc aaa ttt	1244
Val Pro Val Gly Arg Gln Pro Ile Ile Met Gln Ser Ser Ser Lys Phe	
325 330 335	
aac ttt cca tca ggg aga cct gga atg cag aat ggt act gga caa act	1292
Asn Phe Pro Ser Gly Arg Pro Gly Met Gln Asn Gly Thr Gly Gln Thr	
340 345 350	
gat ttc atg ata cac caa aat gtt gtc cct gct ggc act gtg aat cgg	1340
Asp Phe Met Ile His Gln Asn Val Val Pro Ala Gly Thr Val Asn Arg	
355 360 365 370	
cag cca cca cct cca tat cct ctg aca gca gct aat gga caa agc cct	1388
Gln Pro Pro Pro Pro Tyr Pro Leu Thr Ala Ala Asn Gly Gln Ser Pro	
375 380 385	
tct gct tta caa aca ggg gga tct gct gct cct tcg tca tat aca aat	1436
Ser Ala Leu Gln Thr Gly Gly Ser Ala Ala Pro Ser Ser Tyr Thr Asn	
390 395 400	
gga agt att cct cag tct atg atg gtg cca aac aga aat agt cat aac	1484
Gly Ser Ile Pro Gln Ser Met Met Val Pro Asn Arg Asn Ser His Asn	
405 410 415	
atg gaa cta tat aac att agt gta cct gga ctg caa aca aat tgg cct	1532
Met Glu Leu Tyr Asn Ile Ser Val Pro Gly Leu Gln Thr Asn Trp Pro	
420 425 430	
cag tca tct tct gct cca gcc cag tca tcc ccg agc agt ggg cat gaa	1580
Gln Ser Ser Ser Ala Pro Ala Gln Ser Ser Pro Ser Ser Gly His Glu	
435 440 445 450	
atc cct aca tgg caa cct aac ata cca gtg agg tca aat tct ttt aat	1628
Ile Pro Thr Trp Gln Pro Asn Ile Pro Val Arg Ser Asn Ser Phe Asn	
455 460 465	
aac cca tta gga aat aga gca agt cac tct gct aat tct cag cct tct	1676
Asn Pro Leu Gly Asn Arg Ala Ser His Ser Ala Asn Ser Gln Pro Ser	
470 475 480	
gct aca aca gtc act gca att aca cca gct cct att caa cag cct gtg	1724
Ala Thr Thr Val Thr Ala Ile Thr Pro Ala Pro Ile Gln Gln Pro Val	
485 490 495	
aaa agt atg cgt gta tta aaa cca gag cta cag act gct tta gca cct	1772
Lys Ser Met Arg Val Leu Lys Pro Glu Leu Gln Thr Ala Leu Ala Pro	
500 505 510	
aca cac cct tct tgg ata cca cag cca att caa act gtt caa ccc agt	1820
Thr His Pro Ser Trp Ile Pro Gln Pro Ile Gln Thr Val Gln Pro Ser	
515 520 525 530	

cct ttt cct gag gga acc gct tca aat gtg act gtg atg cca cct gtt	1868
Pro Phe Pro Glu Gly Thr Ala Ser Asn Val Thr Val Met Pro Pro Val	
535 540 545	
gct gaa gct cca aac tat caa gga cca cca cca ccc tac cca aaa cat	1916
Ala Glu Ala Pro Asn Tyr Gln Gly Pro Pro Pro Pro Tyr Pro Lys His	
550 555 560	
ctg ctg cac caa aac cca tct gtt cct cca tac gag tca atc agt aag	1964
Leu Leu His Gln Asn Pro Ser Val Pro Pro Tyr Glu Ser Ile Ser Lys	
565 570 575	
cct agc aaa gag gat cag cca agc ttg ccc aag gaa gat gag agt gaa	2012
Pro Ser Lys Glu Asp Gln Pro Ser Leu Pro Lys Glu Asp Glu Ser Glu	
580 585 590	
aag agt tat gaa aat gtt gat agt ggg gat aaa gaa aag aaa cag att	2060
Lys Ser Tyr Glu Asn Val Asp Ser Gly Asp Lys Glu Lys Lys Gln Ile	
595 600 605 610	
aca act tca cct att act gtt agg aaa aac aag aaa gat gaa gag cga	2108
Thr Thr Ser Pro Ile Thr Val Arg Lys Asn Lys Lys Asp Glu Glu Arg	
615 620 625	
agg gaa tct cgt att caa agt tat tct cct caa gca ttt aaa ttc ttt	2156
Arg Glu Ser Arg Ile Gln Ser Tyr Ser Pro Gln Ala Phe Lys Phe Phe	
630 635 640	
atg gag caa cat gta gaa aat gta ctc aaa tct cat cag cag cgt cta	2204
Met Glu Gln His Val Glu Asn Val Leu Lys Ser His Gln Gln Arg Leu	
645 650 655	
cat cgt aaa aaa caa tta gag aat gaa atg atg cgg gtt gga tta tct	2252
His Arg Lys Lys Gln Leu Glu Asn Glu Met Met Arg Val Gly Leu Ser	
660 665 670	
caa gat gcc cag gat caa atg aga aag atg ctt tgc caa aaa gaa tct	2300
Gln Asp Ala Gln Asp Gln Met Arg Lys Met Leu Cys Gln Lys Glu Ser	
675 680 685 690	
aat tac atc cgt ctt aaa agg gct aaa atg gac aag tct atg ttt gtg	2348
Asn Tyr Ile Arg Leu Lys Arg Ala Lys Met Asp Lys Ser Met Phe Val	
695 700 705	
aag ata aag aca cta gga ata gga gca ttt ggt gaa gtc tgt cta gca	2396
Lys Ile Lys Thr Leu Gly Ile Gly Ala Phe Gly Glu Val Cys Leu Ala	
710 715 720	
aga aaa gta gat act aag gct ttg tat gca aca aaa act ctt cga aag	2444
Arg Lys Val Asp Thr Lys Ala Leu Tyr Ala Thr Lys Thr Leu Arg Lys	
725 730 735	
aaa gat gtt ctt ctt cga aat caa gtc gct cat gtt aag gct gag aga	2492
Lys Asp Val Leu Leu Arg Asn Gln Val Ala His Val Lys Ala Glu Arg	
740 745 750	

gat atc ctg gct gaa gct gac aat gaa tgg gta gtt cgt cta tat tat	2540
Asp Ile Leu Ala Glu Ala Asp Asn Glu Trp Val Val Arg Leu Tyr Tyr	
755 760 765 770	
tca ttc caa gat aag gac aat tta tac ttt gta atg gac tac att cct	2588
Ser Phe Gln Asp Lys Asp Asn Leu Tyr Phe Val Met Asp Tyr Ile Pro	
775 780 785	
ggg ggt gat atg atg agc cta tta att aga atg ggc atc ttt cca gaa	2636
Gly Gly Asp Met Met Ser Leu Leu Ile Arg Met Gly Ile Phe Pro Glu	
790 795 800	
agt ctg gca cga ttc tac ata gca gaa ctt acc tgt gca gtt gaa agt	2684
Ser Leu Ala Arg Phe Tyr Ile Ala Glu Leu Thr Cys Ala Val Glu Ser	
805 810 815	
gtt cat aaa atg ggt ttt att cat aga gat att aaa cct gat aat att	2732
Val His Lys Met Gly Phe Ile His Arg Asp Ile Lys Pro Asp Asn Ile	
820 825 830	
ttg att gat cgt gat ggt cat att aaa ttg act gac ttt ggc ctc tgc	2780
Leu Ile Asp Arg Asp Gly His Ile Lys Leu Thr Asp Phe Gly Leu Cys	
835 840 845 850	
act ggc ttc aga tgg aca cac gat tct aag tac tat cag agt ggt gac	2828
Thr Gly Phe Arg Trp Thr His Asp Ser Lys Tyr Tyr Gln Ser Gly Asp	
855 860 865	
cat cca cgg caa gat agc atg gat ttc agt aat gaa tgg ggg gat ccc	2876
His Pro Arg Gln Asp Ser Met Asp Phe Ser Asn Glu Trp Gly Asp Pro	
870 875 880	
tca agc tgt cga tgt gga gac aga ctg aag cca tta gag cgg aga gct	2924
Ser Ser Cys Arg Cys Gly Asp Arg Leu Lys Pro Leu Glu Arg Arg Ala	
885 890 895	
gca cgc cag cac cag cga tgt cta gca cat tct ttg gtt ggg act ccc	2972
Ala Arg Gln His Gln Arg Cys Leu Ala His Ser Leu Val Gly Thr Pro	
900 905 910	
aat tat att gca cct gaa gtg ttg cta cga aca gga tac aca cag ttg	3020
Asn Tyr Ile Ala Pro Glu Val Leu Leu Arg Thr Gly Tyr Thr Gln Leu	
915 920 925 930	
tgt gat tgg tgg agt gtt ggt gtt att ctt ttt gaa atg ttg gtg gga	3068
Cys Asp Trp Trp Ser Val Gly Val Ile Leu Phe Glu Met Leu Val Gly	
935 940 945	
caa cct cct ttc ttg gca caa aca cca tta gaa aca caa atg aag gtt	3116
Gln Pro Pro Phe Leu Ala Gln Thr Pro Leu Glu Thr Gln Met Lys Val	
950 955 960	
atc aac tgg caa aca tct ctt cac att cca cca caa gct aaa ctc agt	3164
Ile Asn Trp Gln Thr Ser Leu His Ile Pro Pro Gln Ala Lys Leu Ser	
965 970 975	

```

cct gaa gct tct gat ctt att att aaa ctt tgc cga gga ccc gaa gat 3212
Pro Glu Ala Ser Asp Leu Ile Ile Lys Leu Cys Arg Gly Pro Glu Asp
980 985 990

cgc tta ggc aag aat ggt gct gat gaa ata aaa gct cat cca ttt ttt 3260
Arg Leu Gly Lys Asn Gly Ala Asp Glu Ile Lys Ala His Pro Phe Phe
995 1000 1005 1010

aaa aca att gac ttc tcc agt gac ctg aga cag cag tct gct tca tac 3308
Lys Thr Ile Asp Phe Ser Ser Asp Leu Arg Gln Gln Ser Ala Ser Tyr
1015 1020 1025

att cct aaa atc aca cac cca aca gat aca tca aat ttt gat cct gtt 3356
Ile Pro Lys Ile Thr His Pro Thr Asp Thr Ser Asn Phe Asp Pro Val
1030 1035 1040

gat cct gat aaa tta tgg agt gat gat aac gag gaa gaa aat gta aat 3404
Asp Pro Asp Lys Leu Trp Ser Asp Asp Asn Glu Glu Glu Asn Val Asn
1045 1050 1055

gac act ctc aat gga tgg tat aaa aat gga aag cat cct gaa cat gca 3452
Asp Thr Leu Asn Gly Trp Tyr Lys Asn Gly Lys His Pro Glu His Ala
1060 1065 1070

ttc tat gaa ttt acc ttc cga agg ttt ttt gat gac aat ggc tac cca 3500
Phe Tyr Glu Phe Thr Phe Arg Arg Phe Phe Asp Asp Asn Gly Tyr Pro
1075 1080 1085 1090

tat aat tat ccg aag cct att gaa tat gaa tac att aat tca caa ggc 3548
Tyr Asn Tyr Pro Lys Pro Ile Glu Tyr Glu Tyr Ile Asn Ser Gln Gly
1095 1100 1105

tca gag cag cag tcg gat gaa gat gat caa aac aca ggc tca gag att 3596
Ser Glu Gln Gln Ser Asp Glu Asp Asp Gln Asn Thr Gly Ser Glu Ile
1110 1115 1120

aaa aat cgc gat cta gta tat gtt taacacacta gtaaataaat gtaatgagga 3650
Lys Asn Arg Asp Leu Val Tyr Val
1125 1130

tttgtaaaaag ggccctgaaat gcgagggtgtt ttgagggttct gagagtaaaa ttatgcaaat 3710

atgacagagc tatatatgtg tgctctgtgt acaatatattt attttcctaa attatgggaa 3770

atccttttaa aatgttaatt tattccagcc gtttaaatca gtatttagaa aaaaattggt 3830

ataaggaaag taaattatga actgaatatt atagtcagtt cttggtactt aaagtactta 3890

aaataagtag tgctttgttt aaaaggagaa acctgggtatc tatttgtata tatgctaaat 3950

aattttaaaa tacaagagtt ttgaaattt tttt 3984

```

&lt;210&gt; 2

&lt;211&gt; 1130

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens



&lt;400&gt; 2

```

Met Lys Arg Ser Glu Lys Pro Glu Gly Tyr Arg Gln Met Arg Pro Lys
 1           5           10           15

Thr Phe Pro Ala Ser Asn Tyr Thr Val Ser Ser Arg Gln Met Leu Gln
          20           25           30

Glu Ile Arg Glu Ser Leu Arg Asn Leu Ser Lys Pro Ser Asp Ala Ala
          35           40           45

Lys Ala Glu His Asn Met Ser Lys Met Ser Thr Glu Asp Pro Arg Gln
          50           55           60

Val Arg Asn Pro Pro Lys Phe Gly Thr His His Lys Ala Leu Gln Glu
          65           70           75           80

Ile Arg Asn Ser Leu Leu Pro Phe Ala Asn Glu Thr Asn Ser Ser Arg
          85           90           95

Ser Thr Ser Glu Val Asn Pro Gln Met Leu Gln Asp Leu Gln Ala Ala
          100          105          110

Gly Phe Asp Glu Asp Met Val Ile Gln Ala Leu Gln Lys Thr Asn Asn
          115          120          125

Arg Ser Ile Glu Ala Ala Ile Glu Phe Ile Ser Lys Met Ser Tyr Gln
          130          135          140

Asp Pro Arg Arg Glu Gln Met Ala Ala Ala Ala Arg Pro Ile Asn
          145          150          155          160

Ala Ser Met Lys Pro Gly Asn Val Gln Gln Ser Val Asn Arg Lys Gln
          165          170          175

Ser Trp Lys Gly Ser Lys Glu Ser Leu Val Pro Gln Arg His Gly Pro
          180          185          190

Pro Leu Gly Glu Ser Val Ala Tyr His Ser Glu Ser Pro Asn Ser Gln
          195          200          205

Thr Asp Val Gly Arg Pro Leu Ser Gly Ser Gly Ile Ser Ala Phe Val
          210          215          220

Gln Ala His Pro Ser Asn Gly Gln Arg Val Asn Pro Pro Pro Pro Pro
          225          230          235          240

Gln Val Arg Ser Val Thr Pro Pro Pro Pro Pro Arg Gly Gln Thr Pro
          245          250          255

Pro Pro Arg Gly Thr Thr Pro Pro Pro Pro Ser Trp Glu Pro Asn Ser
          260          265          270

Gln Thr Lys Arg Tyr Ser Gly Asn Met Glu Tyr Val Ile Ser Arg Ile
          275          280          285

Ser Pro Val Pro Pro Gly Ala Trp Gln Glu Gly Tyr Pro Pro Pro Pro
          290          295          300

```

Leu Asn Thr Ser Pro Met Asn Pro Pro Asn Gln Gly Gln Arg Gly Ile  
 305 310 315 320  
 Ser Ser Val Pro Val Gly Arg Gln Pro Ile Ile Met Gln Ser Ser Ser  
 325 330 335  
 Lys Phe Asn Phe Pro Ser Gly Arg Pro Gly Met Gln Asn Gly Thr Gly  
 340 345 350  
 Gln Thr Asp Phe Met Ile His Gln Asn Val Val Pro Ala Gly Thr Val  
 355 360 365  
 Asn Arg Gln Pro Pro Pro Pro Tyr Pro Leu Thr Ala Ala Asn Gly Gln  
 370 375 380  
 Ser Pro Ser Ala Leu Gln Thr Gly Gly Ser Ala Ala Pro Ser Ser Tyr  
 385 390 395 400  
 Thr Asn Gly Ser Ile Pro Gln Ser Met Met Val Pro Asn Arg Asn Ser  
 405 410 415  
 His Asn Met Glu Leu Tyr Asn Ile Ser Val Pro Gly Leu Gln Thr Asn  
 420 425 430  
 Trp Pro Gln Ser Ser Ser Ala Pro Ala Gln Ser Ser Pro Ser Ser Gly  
 435 440 445  
 His Glu Ile Pro Thr Trp Gln Pro Asn Ile Pro Val Arg Ser Asn Ser  
 450 455 460  
 Phe Asn Asn Pro Leu Gly Asn Arg Ala Ser His Ser Ala Asn Ser Gln  
 465 470 475 480  
 Pro Ser Ala Thr Thr Val Thr Ala Ile Thr Pro Ala Pro Ile Gln Gln  
 485 490 495  
 Pro Val Lys Ser Met Arg Val Leu Lys Pro Glu Leu Gln Thr Ala Leu  
 500 505 510  
 Ala Pro Thr His Pro Ser Trp Ile Pro Gln Pro Ile Gln Thr Val Gln  
 515 520 525  
 Pro Ser Pro Phe Pro Glu Gly Thr Ala Ser Asn Val Thr Val Met Pro  
 530 535 540  
 Pro Val Ala Glu Ala Pro Asn Tyr Gln Gly Pro Pro Pro Pro Tyr Pro  
 545 550 555 560  
 Lys His Leu Leu His Gln Asn Pro Ser Val Pro Pro Tyr Glu Ser Ile  
 565 570 575  
 Ser Lys Pro Ser Lys Glu Asp Gln Pro Ser Leu Pro Lys Glu Asp Glu  
 580 585 590  
 Ser Glu Lys Ser Tyr Glu Asn Val Asp Ser Gly Asp Lys Glu Lys Lys  
 595 600 605

Gln Ile Thr Thr Ser Pro Ile Thr Val Arg Lys Asn Lys Lys Asp Glu  
 610 615 620  
 Glu Arg Arg Glu Ser Arg Ile Gln Ser Tyr Ser Pro Gln Ala Phe Lys  
 625 630 635 640  
 Phe Phe Met Glu Gln His Val Glu Asn Val Leu Lys Ser His Gln Gln  
 645 650 655  
 Arg Leu His Arg Lys Lys Gln Leu Glu Asn Glu Met Met Arg Val Gly  
 660 665 670  
 Leu Ser Gln Asp Ala Gln Asp Gln Met Arg Lys Met Leu Cys Gln Lys  
 675 680 685  
 Glu Ser Asn Tyr Ile Arg Leu Lys Arg Ala Lys Met Asp Lys Ser Met  
 690 695 700  
 Phe Val Lys Ile Lys Thr Leu Gly Ile Gly Ala Phe Gly Glu Val Cys  
 705 710 715 720  
 Leu Ala Arg Lys Val Asp Thr Lys Ala Leu Tyr Ala Thr Lys Thr Leu  
 725 730 735  
 Arg Lys Lys Asp Val Leu Leu Arg Asn Gln Val Ala His Val Lys Ala  
 740 745 750  
 Glu Arg Asp Ile Leu Ala Glu Ala Asp Asn Glu Trp Val Val Arg Leu  
 755 760 765  
 Tyr Tyr Ser Phe Gln Asp Lys Asp Asn Leu Tyr Phe Val Met Asp Tyr  
 770 775 780  
 Ile Pro Gly Gly Asp Met Met Ser Leu Leu Ile Arg Met Gly Ile Phe  
 785 790 795 800  
 Pro Glu Ser Leu Ala Arg Phe Tyr Ile Ala Glu Leu Thr Cys Ala Val  
 805 810 815  
 Glu Ser Val His Lys Met Gly Phe Ile His Arg Asp Ile Lys Pro Asp  
 820 825 830  
 Asn Ile Leu Ile Asp Arg Asp Gly His Ile Lys Leu Thr Asp Phe Gly  
 835 840 845  
 Leu Cys Thr Gly Phe Arg Trp Thr His Asp Ser Lys Tyr Tyr Gln Ser  
 850 855 860  
 Gly Asp His Pro Arg Gln Asp Ser Met Asp Phe Ser Asn Glu Trp Gly  
 865 870 875 880  
 Asp Pro Ser Ser Cys Arg Cys Gly Asp Arg Leu Lys Pro Leu Glu Arg  
 885 890 895  
 Arg Ala Ala Arg Gln His Gln Arg Cys Leu Ala His Ser Leu Val Gly  
 900 905 910

Thr Pro Asn Tyr Ile Ala Pro Glu Val Leu Leu Arg Thr Gly Tyr Thr  
 915 920 925  
 Gln Leu Cys Asp Trp Trp Ser Val Gly Val Ile Leu Phe Glu Met Leu  
 930 935 940  
 Val Gly Gln Pro Pro Phe Leu Ala Gln Thr Pro Leu Glu Thr Gln Met  
 945 950 955 960  
 Lys Val Ile Asn Trp Gln Thr Ser Leu His Ile Pro Pro Gln Ala Lys  
 965 970 975  
 Leu Ser Pro Glu Ala Ser Asp Leu Ile Ile Lys Leu Cys Arg Gly Pro  
 980 985 990  
 Glu Asp Arg Leu Gly Lys Asn Gly Ala Asp Glu Ile Lys Ala His Pro  
 995 1000 1005  
 Phe Phe Lys Thr Ile Asp Phe Ser Ser Asp Leu Arg Gln Gln Ser Ala  
 1010 1015 1020  
 Ser Tyr Ile Pro Lys Ile Thr His Pro Thr Asp Thr Ser Asn Phe Asp  
 025 1030 1035 1040  
 Pro Val Asp Pro Asp Lys Leu Trp Ser Asp Asp Asn Glu Glu Glu Asn  
 1045 1050 1055  
 Val Asn Asp Thr Leu Asn Gly Trp Tyr Lys Asn Gly Lys His Pro Glu  
 1060 1065 1070  
 His Ala Phe Tyr Glu Phe Thr Phe Arg Arg Phe Phe Asp Asp Asn Gly  
 1075 1080 1085  
 Tyr Pro Tyr Asn Tyr Pro Lys Pro Ile Glu Tyr Glu Tyr Ile Asn Ser  
 1090 1095 1100  
 Gln Gly Ser Glu Gln Gln Ser Asp Glu Asp Asp Gln Asn Thr Gly Ser  
 105 1110 1115 1120  
 Glu Ile Lys Asn Arg Asp Leu Val Tyr Val  
 1125 1130

<210> 3  
 <211> 3213  
 <212> DNA  
 <213> Mus musculus

<220>  
 <221> CDS  
 <222> (1)..(2886)

<400> 3  
 gtg caa cat tca att aac cga aaa caa agc tgg aaa ggt tct aaa gag 48  
 Val Gln His Ser Ile Asn Arg Lys Gln Ser Trp Lys Gly Ser Lys Glu  
 1 5 10 15

tct cta gtt cct cag aga cac ggc cca tct cta gga gaa aat gtg gtt	96
Ser Leu Val Pro Gln Arg His Gly Pro Ser Leu Gly Glu Asn Val Val	
20 25 30	
tat cgt tct gaa agc ccc aac tca cag gcg gat gta gga aga cct ctg	144
Tyr Arg Ser Glu Ser Pro Asn Ser Gln Ala Asp Val Gly Arg Pro Leu	
35 40 45	
tct gga tcc ggc att gca gca ttt gct caa gct cac cca agc aat gga	192
Ser Gly Ser Ser Gly Ile Ala Ala Phe Ala Gln Ala His Pro Ser Asn Gly	
50 55 60	
cag aga gtg aac ccc cca cca cca cct caa gtt agg agt gtt act cct	240
Gln Arg Val Asn Pro Pro Pro Pro Pro Gln Val Arg Ser Val Thr Pro	
65 70 75 80	
cca cca cct ccg aga ggc cag acc cca cct ccc cga ggc acc act ccc	288
Pro Pro Pro Pro Arg Gly Gln Thr Pro Pro Pro Arg Gly Thr Thr Pro	
85 90 95	
cct ccc ccc tca tgg gaa cca agc tct cag aca aag cgc tac tct ggg	336
Pro Pro Pro Ser Trp Glu Pro Ser Ser Gln Thr Lys Arg Tyr Ser Gly	
100 105 110	
aac atg gag tac gta atc tcc cga atc tcc cct gtt cca cct ggg gcg	384
Asn Met Glu Tyr Val Ile Ser Arg Ile Ser Pro Val Pro Pro Gly Ala	
115 120 125	
tgg cag gag ggg tac cct cca cca cct ctt acc act tct ccc atg aat	432
Trp Gln Glu Gly Tyr Pro Pro Pro Pro Leu Thr Thr Ser Pro Met Asn	
130 135 140	
ccc cct agc cag gct cag agg gcc att agt tct gtt cca gtt ggt aga	480
Pro Pro Ser Gln Ala Gln Arg Ala Ile Ser Ser Val Pro Val Gly Arg	
145 150 155 160	
caa ccc atc atc atg cag agt act agc aaa ttt aac ttt aca cca ggg	528
Gln Pro Ile Ile Met Gln Ser Thr Ser Lys Phe Asn Phe Thr Pro Gly	
165 170 175	
cga cct gga gtt cag aat ggt ggt ggt cag tct gat ttt atc gtg cac	576
Arg Pro Gly Val Gln Asn Gly Gly Gly Gln Ser Asp Phe Ile Val His	
180 185 190	
caa aat gtc ccc act ggt tct gtg act cgg cag cca cca cct cca tat	624
Gln Asn Val Pro Thr Gly Ser Val Thr Arg Gln Pro Pro Pro Tyr	
195 200 205	
cct ctg acc cca gct aat gga caa agc ccc tct gct tta caa aca ggg	672
Pro Leu Thr Pro Ala Asn Gly Gln Ser Pro Ser Ala Leu Gln Thr Gly	
210 215 220	
gct tct gct gct cca cca tca ttc gcc aat gga aac gtt cct cag tcg	720
Ala Ser Ala Ala Pro Pro Ser Phe Ala Asn Gly Asn Val Pro Gln Ser	
225 230 235 240	

atg atg gtg ccc aac agg aac agt cat aac atg gag ctt tat aat att	768
Met Met Val Pro Asn Arg Asn Ser His Asn Met Glu Leu Tyr Asn Ile	
245 250 255	
aat gtc cct gga ctg caa aca gcc tgg ccc cag tcg tct tct gct cct	816
Asn Val Pro Gly Leu Gln Thr Ala Trp Pro Gln Ser Ser Ser Ala Pro	
260 265 270	
gcg cag tca tcc cca agc ggt ggg cat gaa att cct aca tgg caa cct	864
Ala Gln Ser Ser Pro Ser Gly Gly His Glu Ile Pro Thr Trp Gln Pro	
275 280 285	
aac ata cca gtg agg tca aat tct ttt aat aac cca tta gga agt aga	912
Asn Ile Pro Val Arg Ser Asn Ser Phe Asn Asn Pro Leu Gly Ser Arg	
290 295 300	
gca agt cac tct gct aat tct cag cct tct gcc act aca gtc act gcc	960
Ala Ser His Ser Ala Asn Ser Gln Pro Ser Ala Thr Thr Val Thr Ala	
305 310 315 320	
atc aca ccc gct cct att caa cag ccc gtg aaa agc atg cgc gtc ctg	1008
Ile Thr Pro Ala Pro Ile Gln Gln Pro Val Lys Ser Met Arg Val Leu	
325 330 335	
aaa cca gag ctg cag act gct tta gcc cca acc cat cct tct tgg atg	1056
Lys Pro Glu Leu Gln Thr Ala Leu Ala Pro Thr His Pro Ser Trp Met	
340 345 350	
cca cag cca gtt cag act gtt cag cct acc cct ttt tct gag ggt aca	1104
Pro Gln Pro Val Gln Thr Val Gln Pro Thr Pro Phe Ser Glu Gly Thr	
355 360 365	
gct tca agt gtg cct gtc atc cca cct gtt gct gaa gct cca agc tat	1152
Ala Ser Ser Val Pro Val Ile Pro Pro Val Ala Glu Ala Pro Ser Tyr	
370 375 380	
caa ggt cca cca ccg cct tat cca aaa cat ctg cta cac caa aac cca	1200
Gln Gly Pro Pro Pro Tyr Pro Lys His Leu Leu His Gln Asn Pro	
385 390 395 400	
tct gtc cct cca tat gag tca gta agt aag ccc tgc aaa gat gaa cag	1248
Ser Val Pro Pro Tyr Glu Ser Val Ser Lys Pro Cys Lys Asp Glu Gln	
405 410 415	
cct agc tta ccc aag gaa gat gat agt gag aag agt gcg gac agt ggt	1296
Pro Ser Leu Pro Lys Glu Asp Asp Ser Glu Lys Ser Ala Asp Ser Gly	
420 425 430	
gac tct ggg gat aaa gaa aag aaa cag att aca act tca cct atc act	1344
Asp Ser Gly Asp Lys Glu Lys Lys Gln Ile Thr Thr Ser Pro Ile Thr	
435 440 445	
gtt cgg aaa aac aag aaa gat gaa gaa cga aga gag tct cgg att cag	1392
Val Arg Lys Asn Lys Lys Asp Glu Glu Arg Arg Glu Ser Arg Ile Gln	
450 455 460	

agt tac tcc cca cag gcc ttt aag ttc ttc atg gag cag cac gta gag	1440
Ser Tyr Ser Pro Gln Ala Phe Lys Phe Phe Met Glu Gln His Val Glu	
465 470 475 480	
aac gtc ctg aag tct cat cag cag cgt ctg cat cgg aag aag cag cta	1488
Asn Val Leu Lys Ser His Gln Gln Arg Leu His Arg Lys Lys Gln Leu	
485 490 495	
gaa aat gaa atg atg cgg gtt gga tta tct caa gat gcc cag gat caa	1536
Glu Asn Glu Met Met Arg Val Gly Leu Ser Gln Asp Ala Gln Asp Gln	
500 505 510	
atg aga aag atg ctt tgc cag aaa gag tct aac tat att cgt ctt aaa	1584
Met Arg Lys Met Leu Cys Gln Lys Glu Ser Asn Tyr Ile Arg Leu Lys	
515 520 525	
agg gct aaa atg gac aag tct atg ttt gta aag ata aag aca tta gga	1632
Arg Ala Lys Met Asp Lys Ser Met Phe Val Lys Ile Lys Thr Leu Gly	
530 535 540	
ata gga gcg ttt ggt gaa gtc tgt cta gca aga aaa gtc gat act aaa	1680
Ile Gly Ala Phe Gly Glu Val Cys Leu Ala Arg Lys Val Asp Thr Lys	
545 550 555 560	
gct ttg tat gca aca aag act ctt cga aag aaa gac gtt ctg ctc cga	1728
Ala Leu Tyr Ala Thr Lys Thr Leu Arg Lys Lys Asp Val Leu Leu Arg	
565 570 575	
aat cag gtg gct cat gtg aaa gcg gag agg gat atc cta gca gaa gcc	1776
Asn Gln Val Ala His Val Lys Ala Glu Arg Asp Ile Leu Ala Glu Ala	
580 585 590	
gac aat gag tgg gtg gtc cgc ctg tac tac tct ttc cag gac aag gac	1824
Asp Asn Glu Trp Val Val Arg Leu Tyr Tyr Ser Phe Gln Asp Lys Asp	
595 600 605	
aac ttg tac ttt gtg atg gac tac att cct ggg ggg gat atg atg agc	1872
Asn Leu Tyr Phe Val Met Asp Tyr Ile Pro Gly Gly Asp Met Met Ser	
610 615 620	
cta tta att aga atg ggc atc ttt cct gaa aat ctg gca cga ttc tac	1920
Leu Leu Ile Arg Met Gly Ile Phe Pro Glu Asn Leu Ala Arg Phe Tyr	
625 630 635 640	
ata gca gaa ctt acc tgt gca gtt gaa agt gtt cat aaa atg ggt ttt	1968
Ile Ala Glu Leu Thr Cys Ala Val Glu Ser Val His Lys Met Gly Phe	
645 650 655	
att cat aga gat att aaa cct gat aac att ttg att gac cgt gat ggc	2016
Ile His Arg Asp Ile Lys Pro Asp Asn Ile Leu Ile Asp Arg Asp Gly	
660 665 670	
cat att aaa ttg act gac ttt ggc ttg tgc act ggc ttc aga tgg aca	2064
His Ile Lys Leu Thr Asp Phe Gly Leu Cys Thr Gly Phe Arg Trp Thr	
675 680 685	

cat gac tcc aag tac tac cag agt ggg gat cac cca cgg caa gat agc	2112
His Asp Ser Lys Tyr Tyr Gln Ser Gly Asp His Pro Arg Gln Asp Ser	
690 695 700	
atg gat ttc agt aac gaa tgg gga gat cct tcc aat tgt cgg tgt ggg	2160
Met Asp Phe Ser Asn Glu Trp Gly Asp Pro Ser Asn Cys Arg Cys Gly	
705 710 715 720	
gac aga ctg aag cca ctg gag cgg aga gct gct cgc cag cac cag cga	2208
Asp Arg Leu Lys Pro Leu Glu Arg Arg Ala Ala Arg Gln His Gln Arg	
725 730 735	
tgt cta gcc cat tct ctg gtt ggg act ccc aat tat att gca cct gaa	2256
Cys Leu Ala His Ser Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu	
740 745 750	
gtg cta ctg cga aca gga tat aca cag ctg tgt gac tgg tgg agt gtt	2304
Val Leu Leu Arg Thr Gly Tyr Thr Gln Leu Cys Asp Trp Trp Ser Val	
755 760 765	
ggg gtt att ctt tgt gaa atg ttg gtg gga caa cct cct ttc ttg gca	2352
Gly Val Ile Leu Cys Glu Met Leu Val Gly Gln Pro Pro Phe Leu Ala	
770 775 780	
caa acc cca tta gaa aca caa atg aag gtt atc atc tgg caa act tct	2400
Gln Thr Pro Leu Glu Thr Gln Met Lys Val Ile Ile Trp Gln Thr Ser	
785 790 795 800	
cta cac atc cct cct caa gct aag ctg agt cct gaa gcc tct gac ctc	2448
Leu His Ile Pro Pro Gln Ala Lys Leu Ser Pro Glu Ala Ser Asp Leu	
805 810 815	
att atc aaa ctg tgt cga gga cca gaa gac cgc ctc ggc aag aac ggt	2496
Ile Ile Lys Leu Cys Arg Gly Pro Glu Asp Arg Leu Gly Lys Asn Gly	
820 825 830	
gct gat gag ata aag gct cat cca ttt ttt aag acc atc gat ttc tct	2544
Ala Asp Glu Ile Lys Ala His Pro Phe Phe Lys Thr Ile Asp Phe Ser	
835 840 845	
agt gat ctg aga cag cag tct gct tca tac atc cct aaa atc acg cat	2592
Ser Asp Leu Arg Gln Gln Ser Ala Ser Tyr Ile Pro Lys Ile Thr His	
850 855 860	
cca aca gat aca tcc aat ttc gac cct gtt gat cct gat aaa ttg tgg	2640
Pro Thr Asp Thr Ser Asn Phe Asp Pro Val Asp Pro Asp Lys Leu Trp	
865 870 875 880	
agc gat ggc agc gag gag gaa aat atc agt gac act ctg agc gga tgg	2688
Ser Asp Gly Ser Glu Glu Glu Asn Ile Ser Asp Thr Leu Ser Gly Trp	
885 890 895	
tat aaa aat ggg aag cac ccc gag cac gct ttc tat gag ttc acc ttt	2736
Tyr Lys Asn Gly Lys His Pro Glu His Ala Phe Tyr Glu Phe Thr Phe	
900 905 910	



cgg agg ttt ttt gat gac aat ggc tac cca tat aat tat cca aag cct 2784  
 Arg Arg Phe Phe Asp Asp Asn Gly Tyr Pro Tyr Asn Tyr Pro Lys Pro  
           915                                  920                                  925

att gag tat gaa tac att cat tca cag ggc tca gaa caa cag tct gat 2832  
 Ile Glu Tyr Glu Tyr Ile His Ser Gln Gly Ser Glu Gln Gln Ser Asp  
           930                                  935                                  940

gaa gat gat caa cac aca agc tcc gat gga aac aac cga gat cta gtg 2880  
 Glu Asp Asp Gln His Thr Ser Ser Asp Gly Asn Asn Arg Asp Leu Val  
           945                                  950                                  955                                  960

tat gtt taataaacta ggagatcatt gtaagaattt gcaagaggcc tgaagtgcag 2936  
 Tyr Val

gggtttttga agttttgaga aaattatgca aatgtgacag agtttgtgtg ctctgtgtac 2996

aatatttttat tttcctaagt tatgggaaat tgttttaaaa tggttaattta ttccaccctt 3056

ttaattcagt aatttagaaa aaattgttat aaggaaagta aattatgaac tgagtattat 3116

agtcaattct tggtacttaa agtacttaaa aagagaagcc tggtatcttt tgtatatata 3176

ataaataatt ttaaaatccc aaaaaaaaaa aaaaaaa 3213

<210> 4

<211> 962

<212> PRT

<213> Mus musculus

<400> 4

Val Gln His Ser Ile Asn Arg Lys Gln Ser Trp Lys Gly Ser Lys Glu  
       1                                  5                                  10                                  15

Ser Leu Val Pro Gln Arg His Gly Pro Ser Leu Gly Glu Asn Val Val  
           20                                  25                                  30

Tyr Arg Ser Glu Ser Pro Asn Ser Gln Ala Asp Val Gly Arg Pro Leu  
           35                                  40                                  45

Ser Gly Ser Gly Ile Ala Ala Phe Ala Gln Ala His Pro Ser Asn Gly  
           50                                  55                                  60

Gln Arg Val Asn Pro Pro Pro Pro Pro Gln Val Arg Ser Val Thr Pro  
           65                                  70                                  75                                  80

Pro Pro Pro Pro Arg Gly Gln Thr Pro Pro Pro Arg Gly Thr Thr Pro  
                                   85                                  90                                  95

Pro Pro Pro Ser Trp Glu Pro Ser Ser Gln Thr Lys Arg Tyr Ser Gly  
           100                                  105                                  110

Asn Met Glu Tyr Val Ile Ser Arg Ile Ser Pro Val Pro Pro Gly Ala  
           115                                  120                                  125

Trp Gln Glu Gly Tyr Pro Pro Pro Pro Leu Thr Thr Ser Pro Met Asn  
           130                                  135                                  140

Pro Pro Ser Gln Ala Gln Arg Ala Ile Ser Ser Val Pro Val Gly Arg  
 145 150 155 160  
 Gln Pro Ile Ile Met Gln Ser Thr Ser Lys Phe Asn Phe Thr Pro Gly  
 165 170 175  
 Arg Pro Gly Val Gln Asn Gly Gly Gly Gln Ser Asp Phe Ile Val His  
 180 185 190  
 Gln Asn Val Pro Thr Gly Ser Val Thr Arg Gln Pro Pro Pro Pro Tyr  
 195 200 205  
 Pro Leu Thr Pro Ala Asn Gly Gln Ser Pro Ser Ala Leu Gln Thr Gly  
 210 215 220  
 Ala Ser Ala Ala Pro Pro Ser Phe Ala Asn Gly Asn Val Pro Gln Ser  
 225 230 235 240  
 Met Met Val Pro Asn Arg Asn Ser His Asn Met Glu Leu Tyr Asn Ile  
 245 250 255  
 Asn Val Pro Gly Leu Gln Thr Ala Trp Pro Gln Ser Ser Ser Ala Pro  
 260 265 270  
 Ala Gln Ser Ser Pro Ser Gly Gly His Glu Ile Pro Thr Trp Gln Pro  
 275 280 285  
 Asn Ile Pro Val Arg Ser Asn Ser Phe Asn Asn Pro Leu Gly Ser Arg  
 290 295 300  
 Ala Ser His Ser Ala Asn Ser Gln Pro Ser Ala Thr Thr Val Thr Ala  
 305 310 315 320  
 Ile Thr Pro Ala Pro Ile Gln Gln Pro Val Lys Ser Met Arg Val Leu  
 325 330 335  
 Lys Pro Glu Leu Gln Thr Ala Leu Ala Pro Thr His Pro Ser Trp Met  
 340 345 350  
 Pro Gln Pro Val Gln Thr Val Gln Pro Thr Pro Phe Ser Glu Gly Thr  
 355 360 365  
 Ala Ser Ser Val Pro Val Ile Pro Pro Val Ala Glu Ala Pro Ser Tyr  
 370 375 380  
 Gln Gly Pro Pro Pro Pro Tyr Pro Lys His Leu Leu His Gln Asn Pro  
 385 390 395 400  
 Ser Val Pro Pro Tyr Glu Ser Val Ser Lys Pro Cys Lys Asp Glu Gln  
 405 410 415  
 Pro Ser Leu Pro Lys Glu Asp Asp Ser Glu Lys Ser Ala Asp Ser Gly  
 420 425 430  
 Asp Ser Gly Asp Lys Glu Lys Lys Gln Ile Thr Thr Ser Pro Ile Thr  
 435 440 445

Val Arg Lys Asn Lys Lys Asp Glu Glu Arg Arg Glu Ser Arg Ile Gln  
 450 455 460  
 Ser Tyr Ser Pro Gln Ala Phe Lys Phe Phe Met Glu Gln His Val Glu  
 465 470 475 480  
 Asn Val Leu Lys Ser His Gln Gln Arg Leu His Arg Lys Lys Gln Leu  
 485 490 495  
 Glu Asn Glu Met Met Arg Val Gly Leu Ser Gln Asp Ala Gln Asp Gln  
 500 505 510  
 Met Arg Lys Met Leu Cys Gln Lys Glu Ser Asn Tyr Ile Arg Leu Lys  
 515 520 525  
 Arg Ala Lys Met Asp Lys Ser Met Phe Val Lys Ile Lys Thr Leu Gly  
 530 535 540  
 Ile Gly Ala Phe Gly Glu Val Cys Leu Ala Arg Lys Val Asp Thr Lys  
 545 550 555 560  
 Ala Leu Tyr Ala Thr Lys Thr Leu Arg Lys Lys Asp Val Leu Leu Arg  
 565 570 575  
 Asn Gln Val Ala His Val Lys Ala Glu Arg Asp Ile Leu Ala Glu Ala  
 580 585 590  
 Asp Asn Glu Trp Val Val Arg Leu Tyr Tyr Ser Phe Gln Asp Lys Asp  
 595 600 605  
 Asn Leu Tyr Phe Val Met Asp Tyr Ile Pro Gly Gly Asp Met Met Ser  
 610 615 620  
 Leu Leu Ile Arg Met Gly Ile Phe Pro Glu Asn Leu Ala Arg Phe Tyr  
 625 630 635 640  
 Ile Ala Glu Leu Thr Cys Ala Val Glu Ser Val His Lys Met Gly Phe  
 645 650 655  
 Ile His Arg Asp Ile Lys Pro Asp Asn Ile Leu Ile Asp Arg Asp Gly  
 660 665 670  
 His Ile Lys Leu Thr Asp Phe Gly Leu Cys Thr Gly Phe Arg Trp Thr  
 675 680 685  
 His Asp Ser Lys Tyr Tyr Gln Ser Gly Asp His Pro Arg Gln Asp Ser  
 690 695 700  
 Met Asp Phe Ser Asn Glu Trp Gly Asp Pro Ser Asn Cys Arg Cys Gly  
 705 710 715 720  
 Asp Arg Leu Lys Pro Leu Glu Arg Arg Ala Ala Arg Gln His Gln Arg  
 725 730 735  
 Cys Leu Ala His Ser Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu  
 740 745 750

```

<400> 5
atg aga gcc acc ccg aag ttt gga cct tat caa aaa gct ctc agg gaa    48
Met Arg Ala Thr Pro Lys Phe Gly Pro Tyr Gln Lys Ala Leu Arg Glu
      1              5              10              15

```

atc	cga	tat	tcc	ctc	ctg	cct	ttt	gcc	aac	gag	tca	ggc	act	tcg	gca	96
Ile	Arg	Tyr	Ser	Leu	Leu	Pro	Phe	Ala	Asn	Glu	Ser	Gly	Thr	Ser	Ala	
			20					25					30			
gct	gca	gag	gtg	aac	cgg	cag	atg	ctt	cag	gag	ttg	gtg	aat	gcg	gca	144
Ala	Ala	Glu	Val	Asn	Arg	Gln	Met	Leu	Gln	Glu	Leu	Val	Asn	Ala	Ala	
		35					40					45				
tgt	gac	cag	gag	atg	gct	ggc	aga	gcg	ctc	acg	cag	acg	ggc	agt	agg	192
Cys	Asp	Gln	Glu	Met	Ala	Gly	Arg	Ala	Leu	Thr	Gln	Thr	Gly	Ser	Arg	
	50					55					60					
agt	atc	gaa	gct	gcc	ttg	gag	tac	atc	agt	aag	atg	ggc	tac	ctg	gac	240
Ser	Ile	Glu	Ala	Ala	Leu	Glu	Tyr	Ile	Ser	Lys	Met	Gly	Tyr	Leu	Asp	
	65				70					75					80	
ccc	agg	aat	gag	cag	att	gtg	cga	gtc	atc	aag	cag	acc	tcc	cca	gga	288
Pro	Arg	Asn	Glu	Gln	Ile	Val	Arg	Val	Ile	Lys	Gln	Thr	Ser	Pro	Gly	
				85					90					95		
aag	ggc	ctg	gcg	tcc	acc	ccg	gtg	act	cgg	cgg	ccc	agt	ttc	gag	ggc	336
Lys	Gly	Leu	Ala	Ser	Thr	Pro	Val	Thr	Arg	Arg	Pro	Ser	Phe	Glu	Gly	
		100						105					110			
aca	ggg	gaa	gca	ctc	cca	tcc	tac	cac	cag	ctg	ggt	ggt	gca	aac	tac	384
Thr	Gly	Glu	Ala	Leu	Pro	Ser	Tyr	His	Gln	Leu	Gly	Gly	Ala	Asn	Tyr	
		115					120					125				
gag	ggc	ccc	gcc	gca	ctg	gag	gag	atg	ccg	cgg	caa	tat	tta	gac	ttt	432
Glu	Gly	Pro	Ala	Ala	Leu	Glu	Glu	Met	Pro	Arg	Gln	Tyr	Leu	Asp	Phe	
	130					135					140					
ctc	ttc	cct	gga	gcc	gga	gcc	ggc	acc	cac	ggt	gcc	cag	gct	cac	cag	480
Leu	Phe	Pro	Gly	Ala	Gly	Ala	Gly	Thr	His	Gly	Ala	Gln	Ala	His	Gln	
	145				150					155					160	
cat	cct	ccc	aaa	ggg	tac	agc	aca	gca	gta	gag	cca	agt	gcg	cac	ttt	528
His	Pro	Pro	Lys	Gly	Tyr	Ser	Thr	Ala	Val	Glu	Pro	Ser	Ala	His	Phe	
			165						170					175		
ccg	ggc	aca	cac	tat	ggt	cgt	ggt	cat	cta	cta	tcg	gag	cag	tct	ggg	576
Pro	Gly	Thr	His	Tyr	Gly	Arg	Gly	His	Leu	Leu	Ser	Glu	Gln	Ser	Gly	
			180					185					190			
tat	ggg	gtg	cag	cgc	agt	tcc	tcc	ttc	cag	aac	aag	acg	cca	cca	gat	624
Tyr	Gly	Val	Gln	Arg	Ser	Ser	Ser	Phe	Gln	Asn	Lys	Thr	Pro	Pro	Asp	
		195					200					205				
gcc	tat	tcc	agc	atg	gcc	aag	gcc	cag	ggt	ggc	cct	ccc	gcc	agc	ctc	672
Ala	Tyr	Ser	Ser	Met	Ala	Lys	Ala	Gln	Gly	Gly	Pro	Pro	Ala	Ser	Leu	
	210					215					220					
acc	ttt	cct	gcc	cat	gct	ggg	ctg	tac	act	gcc	tcg	cac	cac	aag	ccg	720
Thr	Phe	Pro	Ala	His	Ala	Gly	Leu	Tyr	Thr	Ala	Ser	His	His	Lys	Pro	
	225				230					235					240	

gcg gct acc cca cct ggg gcc cac cca tta cat gtg ttg ggc acc cgg	768
Ala Ala Thr Pro Pro Gly Ala His Pro Leu His Val Leu Gly Thr Arg	
245 250 255	
ggt ccc acg ttt act ggc gaa agc tct gca cag gct gtg ctg gca ccg	816
Gly Pro Thr Phe Thr Gly Glu Ser Ser Ala Gln Ala Val Leu Ala Pro	
260 265 270	
tcc agg aac agc ctc aat gct gac ttg tac gag ctg ggc tcc acg gtg	864
Ser Arg Asn Ser Leu Asn Ala Asp Leu Tyr Glu Leu Gly Ser Thr Val	
275 280 285	
ccc tgg tct gca gct cca ctg gca cgc cgc gac tcg ctg cag aag cag	912
Pro Trp Ser Ala Ala Pro Leu Ala Arg Arg Asp Ser Leu Gln Lys Gln	
290 295 300	
ggt cta gaa gcc tcg cgg ccg cat gtg gct ttt cgg gct ggc ccc agc	960
Gly Leu Glu Ala Ser Arg Pro His Val Ala Phe Arg Ala Gly Pro Ser	
305 310 315 320	
agg acc aac tcc ttc aac aac cca caa cct gag ccc tca ctg ccc gcc	1008
Arg Thr Asn Ser Phe Asn Asn Pro Gln Pro Glu Pro Ser Leu Pro Ala	
325 330 335	
ccc aac acg gtc acc gcc gtg acg gcc gca cac atc ctt cac cct gtg	1056
Pro Asn Thr Val Thr Ala Val Thr Ala Ala His Ile Leu His Pro Val	
340 345 350	
aag agc gtg cgt gtg ctg cgg ccc gag ccc cag aca gcc gtg ggg ccc	1104
Lys Ser Val Arg Val Leu Arg Pro Glu Pro Gln Thr Ala Val Gly Pro	
355 360 365	
tcg cac ccc gcc tgg gtg gct gcg ccc aca gca cct gcc act gag agc	1152
Ser His Pro Ala Trp Val Ala Ala Pro Thr Ala Pro Ala Thr Glu Ser	
370 375 380	
ctg gag acg aag gag ggc agc gca ggc cca cac ccg ctg gat gtg gac	1200
Leu Glu Thr Lys Glu Gly Ser Ala Gly Pro His Pro Leu Asp Val Asp	
385 390 395 400	
tat ggc ggc tcc gag cgc agg tgc cca ccg cct ccg tat cca aag cac	1248
Tyr Gly Gly Ser Glu Arg Arg Cys Pro Pro Pro Pro Tyr Pro Lys His	
405 410 415	
ttg ctg ctg ccc agt aag tct gag cag tac agc gtg gac ctg gac agc	1296
Leu Leu Leu Pro Ser Lys Ser Glu Gln Tyr Ser Val Asp Leu Asp Ser	
420 425 430	
ctg tgc acc agt gtg cag cag agt ctg cga ggg ggc act gat cta gac	1344
Leu Cys Thr Ser Val Gln Gln Ser Leu Arg Gly Gly Thr Asp Leu Asp	
435 440 445	
ggg agt gac aag agc cac aaa ggt gcg aag gga gac aaa gct ggc aga	1392
Gly Ser Asp Lys Ser His Lys Gly Ala Lys Gly Asp Lys Ala Gly Arg	
450 455 460	

gac	aaa	aag	cag	att	cag	acc	tcc	ccg	gtg	cct	gtc	cgc	aag	aat	agc	1440
Asp	Lys	Lys	Gln	Ile	Gln	Thr	Ser	Pro	Val	Pro	Val	Arg	Lys	Asn	Ser	
465					470					475					480	
aga	gat	gaa	gag	aag	aga	gag	tct	cgc	atc	aag	agt	tac	tcc	cct	tat	1488
Arg	Asp	Glu	Glu	Lys	Arg	Glu	Ser	Arg	Ile	Lys	Ser	Tyr	Ser	Pro	Tyr	
				485					490					495		
gcc	ttc	aaa	ttc	ttc	atg	gag	caa	cac	gtg	gag	aat	gtc	atc	aaa	acc	1536
Ala	Phe	Lys	Phe	Phe	Met	Glu	Gln	His	Val	Glu	Asn	Val	Ile	Lys	Thr	
			500					505					510			
tac	cag	cag	aag	gtc	agc	cgg	agg	cta	cag	ctg	gag	cag	gaa	atg	gcc	1584
Tyr	Gln	Gln	Lys	Val	Ser	Arg	Arg	Leu	Gln	Leu	Glu	Gln	Glu	Met	Ala	
			515				520						525			
aaa	gct	ggg	ctc	tgt	gag	gcc	gag	cag	gag	cag	atg	agg	aag	atc	ctc	1632
Lys	Ala	Gly	Leu	Cys	Glu	Ala	Glu	Gln	Glu	Gln	Met	Arg	Lys	Ile	Leu	
	530					535					540					
tac	cag	aag	gag	tct	aac	tac	aac	cgg	ctg	aag	agg	gcc	aag	atg	gac	1680
Tyr	Gln	Lys	Glu	Ser	Asn	Tyr	Asn	Arg	Leu	Lys	Arg	Ala	Lys	Met	Asp	
545					550					555					560	
aag	tcc	atg	ttt	gtg	aaa	atc	aag	act	cta	ggc	atc	ggg	gcc	ttt	ggg	1728
Lys	Ser	Met	Phe	Val	Lys	Ile	Lys	Thr	Leu	Gly	Ile	Gly	Ala	Phe	Gly	
				565					570					575		
gaa	gtg	tgc	ctc	gct	tgt	aag	ctg	gac	act	cac	gct	ctg	tac	gcc	atg	1776
Glu	Val	Cys	Leu	Ala	Cys	Lys	Leu	Asp	Thr	His	Ala	Leu	Tyr	Ala	Met	
			580					585					590			
aag	act	ctc	agg	aag	aag	gat	gtc	ctg	aac	cgg	aat	caa	gtg	gcc	cat	1824
Lys	Thr	Leu	Arg	Lys	Lys	Asp	Val	Leu	Asn	Arg	Asn	Gln	Val	Ala	His	
			595				600					605				
gtc	aag	gct	gag	agg	gac	atc	ctg	gct	gaa	gca	gac	aat	gag	tgg	gtg	1872
Val	Lys	Ala	Glu	Arg	Asp	Ile	Leu	Ala	Glu	Ala	Asp	Asn	Glu	Trp	Val	
	610					615					620					
gtc	aaa	ctc	tac	tac	tcc	ttc	cag	gac	aag	gac	agc	ctg	tac	ttt	gtg	1920
Val	Lys	Leu	Tyr	Tyr	Ser	Phe	Gln	Asp	Lys	Asp	Ser	Leu	Tyr	Phe	Val	
625					630					635					640	
atg	gac	tac	ata	cca	ggc	ggg	gat	atg	atg	agc	ctg	ctg	atc	agg	atg	1968
Met	Asp	Tyr	Ile	Pro	Gly	Gly	Asp	Met	Met	Ser	Leu	Leu	Ile	Arg	Met	
				645					650					655		
gag	gtc	ttc	cct	gag	cac	ctg	gcc	cgc	ttc	tac	att	gca	gag	ttg	acc	2016
Glu	Val	Phe	Pro	Glu	His	Leu	Ala	Arg	Phe	Tyr	Ile	Ala	Glu	Leu	Thr	
			660					665					670			
ctg	gcc	att	gaa	agt	gtc	cac	aag	atg	ggc	ttt	atc	cac	cgg	gac	atc	2064
Leu	Ala	Ile	Glu	Ser	Val	His	Lys	Met	Gly	Phe	Ile	His	Arg	Asp	Ile	
		675					680					685				

aag cct gac aac ata ctc atc gac ctg gat ggt cat att aag ctg aca Lys Pro Asp Asn Ile Leu Ile Asp Leu Asp Gly His Ile Lys Leu Thr 690 695 700	2112
gat ttt ggc ctc tgc act gga ttc agg tgg act cac aat tcc aag tac Asp Phe Gly Leu Cys Thr Gly Phe Arg Trp Thr His Asn Ser Lys Tyr 705 710 715 720	2160
tac cag aaa ggg aac cac atg aga cag gac agc atg gag ccc ggt gac Tyr Gln Lys Gly Asn His Met Arg Gln Asp Ser Met Glu Pro Gly Asp 725 730 735	2208
ctc tgg gac gat gtt tcc aac tgt cgc tgt gga gac agg tta aag acc Leu Trp Asp Asp Val Ser Asn Cys Arg Cys Gly Asp Arg Leu Lys Thr 740 745 750	2256
ctg gag cag agg gcg cag aag cag cac cag agg tgc ctg gca cat tct Leu Glu Gln Arg Ala Gln Lys Gln His Gln Arg Cys Leu Ala His Ser 755 760 765	2304
ctt gtc ggg aca cca aat tac atc gct ccg gag gtg ctt ctc cgc aaa Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu Val Leu Leu Arg Lys 770 775 780	2352
ggg tac acg cag ctc tgt gac tgg tgg agc gtc ggt gtg att ctc ttt Gly Tyr Thr Gln Leu Cys Asp Trp Trp Ser Val Gly Val Ile Leu Phe 785 790 795 800	2400
gag atg ctg gtt ggg cag ccg cct ttc ttg gcc ccc acc ccc aca gag Glu Met Leu Val Gly Gln Pro Pro Phe Leu Ala Pro Thr Pro Thr Glu 805 810 815	2448
acg cag ctg aag gtg atc aac tgg gag agc acg ctg cat atc cct acg Thr Gln Leu Lys Val Ile Asn Trp Glu Ser Thr Leu His Ile Pro Thr 820 825 830	2496
cag gtg agg ctc agc gct gag gcc cga gac ctc atc acg aag ctg tgc Gln Val Arg Leu Ser Ala Glu Ala Arg Asp Leu Ile Thr Lys Leu Cys 835 840 845	2544
tgc gcg gct gac tgc cgc ctg ggc agg gat ggg gca gat gac ctc aag Cys Ala Ala Asp Cys Arg Leu Gly Arg Asp Gly Ala Asp Asp Leu Lys 850 855 860	2592
gca cac ccg ttc ttc aac acc atc gac ttt tcc cgt gac atc cga aag Ala His Pro Phe Phe Asn Thr Ile Asp Phe Ser Arg Asp Ile Arg Lys 865 870 875 880	2640
cag gct gca ccc tac gtc ccc acc atc agc cac ccc atg gac acc tcc Gln Ala Ala Pro Tyr Val Pro Thr Ile Ser His Pro Met Asp Thr Ser 885 890 895	2688
aat ttt gac ccg gtg gat gaa gaa agc ccc tgg cac gag gcc agc gga Asn Phe Asp Pro Val Asp Glu Glu Ser Pro Trp His Glu Ala Ser Gly 900 905 910	2736



gag agc gcc aag gcc tgg gac acg ctg gcc tcc ccc agc agc aag cat 2784  
 Glu Ser Ala Lys Ala Trp Asp Thr Leu Ala Ser Pro Ser Ser Lys His  
           915                          920                          925

cca gag cac gcc ttc tat gag ttc acc ttc cgc agg ttc ttc gat gac 2832  
 Pro Glu His Ala Phe Tyr Glu Phe Thr Phe Arg Arg Phe Phe Asp Asp  
           930                          935                          940

aac ggc tat ccc ttc cgg tgc ccg aag ccc tca gag ccc gca gag agt 2880  
 Asn Gly Tyr Pro Phe Arg Cys Pro Lys Pro Ser Glu Pro Ala Glu Ser  
           945                          950                          955                          960

gca gac cca ggg gat gcg gac ttg gaa ggt gcg gcc gag ggc tgc cag 2928  
 Ala Asp Pro Gly Asp Ala Asp Leu Glu Gly Ala Ala Glu Gly Cys Gln  
                           965                          970                          975

ccg gtg tac gtg taagcctcag ttaaccacaa ctcgaggaaa cccaaaatga 2980  
 Pro Val Tyr Val  
                           980

gattttctttt cagaagacaa actcaagctt aggaatcctt catttttagt tctggtaaata 3040

gggcaacagg aagagtcaac atgatttcaa attagccctc tgaggacctt cactgcatta 3100

aaacagtatt ttttaaaaaa ttagtacagt atggaaagag cacttatttt gggggg 3155

&lt;210&gt; 6

&lt;211&gt; 980

&lt;212&gt; PRT

&lt;213&gt; Mus musculus

&lt;400&gt; 6

Met Arg Ala Thr Pro Lys Phe Gly Pro Tyr Gln Lys Ala Leu Arg Glu  
   1                          5                          10                          15

Ile Arg Tyr Ser Leu Leu Pro Phe Ala Asn Glu Ser Gly Thr Ser Ala  
           20                          25                          30

Ala Ala Glu Val Asn Arg Gln Met Leu Gln Glu Leu Val Asn Ala Ala  
           35                          40                          45

Cys Asp Gln Glu Met Ala Gly Arg Ala Leu Thr Gln Thr Gly Ser Arg  
   50                          55                          60

Ser Ile Glu Ala Ala Leu Glu Tyr Ile Ser Lys Met Gly Tyr Leu Asp  
   65                          70                          75                          80

Pro Arg Asn Glu Gln Ile Val Arg Val Ile Lys Gln Thr Ser Pro Gly  
           85                          90                          95

Lys Gly Leu Ala Ser Thr Pro Val Thr Arg Arg Pro Ser Phe Glu Gly  
           100                          105                          110

Thr Gly Glu Ala Leu Pro Ser Tyr His Gln Leu Gly Gly Ala Asn Tyr  
           115                          120                          125

Glu Gly Pro Ala Ala Leu Glu Glu Met Pro Arg Gln Tyr Leu Asp Phe  
 130 135 140  
 Leu Phe Pro Gly Ala Gly Ala Gly Thr His Gly Ala Gln Ala His Gln  
 145 150 155 160  
 His Pro Pro Lys Gly Tyr Ser Thr Ala Val Glu Pro Ser Ala His Phe  
 165 170 175  
 Pro Gly Thr His Tyr Gly Arg Gly His Leu Leu Ser Glu Gln Ser Gly  
 180 185 190  
 Tyr Gly Val Gln Arg Ser Ser Ser Phe Gln Asn Lys Thr Pro Pro Asp  
 195 200 205  
 Ala Tyr Ser Ser Met Ala Lys Ala Gln Gly Gly Pro Pro Ala Ser Leu  
 210 215 220  
 Thr Phe Pro Ala His Ala Gly Leu Tyr Thr Ala Ser His His Lys Pro  
 225 230 235 240  
 Ala Ala Thr Pro Pro Gly Ala His Pro Leu His Val Leu Gly Thr Arg  
 245 250 255  
 Gly Pro Thr Phe Thr Gly Glu Ser Ser Ala Gln Ala Val Leu Ala Pro  
 260 265 270  
 Ser Arg Asn Ser Leu Asn Ala Asp Leu Tyr Glu Leu Gly Ser Thr Val  
 275 280 285  
 Pro Trp Ser Ala Ala Pro Leu Ala Arg Arg Asp Ser Leu Gln Lys Gln  
 290 295 300  
 Gly Leu Glu Ala Ser Arg Pro His Val Ala Phe Arg Ala Gly Pro Ser  
 305 310 315 320  
 Arg Thr Asn Ser Phe Asn Asn Pro Gln Pro Glu Pro Ser Leu Pro Ala  
 325 330 335  
 Pro Asn Thr Val Thr Ala Val Thr Ala Ala His Ile Leu His Pro Val  
 340 345 350  
 Lys Ser Val Arg Val Leu Arg Pro Glu Pro Gln Thr Ala Val Gly Pro  
 355 360 365  
 Ser His Pro Ala Trp Val Ala Ala Pro Thr Ala Pro Ala Thr Glu Ser  
 370 375 380  
 Leu Glu Thr Lys Glu Gly Ser Ala Gly Pro His Pro Leu Asp Val Asp  
 385 390 395 400  
 Tyr Gly Gly Ser Glu Arg Arg Cys Pro Pro Pro Tyr Pro Lys His  
 405 410 415  
 Leu Leu Leu Pro Ser Lys Ser Glu Gln Tyr Ser Val Asp Leu Asp Ser  
 420 425 430

Leu Cys Thr Ser Val Gln Gln Ser Leu Arg Gly Gly Thr Asp Leu Asp  
 435 440 445  
 Gly Ser Asp Lys Ser His Lys Gly Ala Lys Gly Asp Lys Ala Gly Arg  
 450 455 460  
 Asp Lys Lys Gln Ile Gln Thr Ser Pro Val Pro Val Arg Lys Asn Ser  
 465 470 475 480  
 Arg Asp Glu Glu Lys Arg Glu Ser Arg Ile Lys Ser Tyr Ser Pro Tyr  
 485 490 495  
 Ala Phe Lys Phe Phe Met Glu Gln His Val Glu Asn Val Ile Lys Thr  
 500 505 510  
 Tyr Gln Gln Lys Val Ser Arg Arg Leu Gln Leu Glu Gln Glu Met Ala  
 515 520 525  
 Lys Ala Gly Leu Cys Glu Ala Glu Gln Glu Gln Met Arg Lys Ile Leu  
 530 535 540  
 Tyr Gln Lys Glu Ser Asn Tyr Asn Arg Leu Lys Arg Ala Lys Met Asp  
 545 550 555 560  
 Lys Ser Met Phe Val Lys Ile Lys Thr Leu Gly Ile Gly Ala Phe Gly  
 565 570 575  
 Glu Val Cys Leu Ala Cys Lys Leu Asp Thr His Ala Leu Tyr Ala Met  
 580 585 590  
 Lys Thr Leu Arg Lys Lys Asp Val Leu Asn Arg Asn Gln Val Ala His  
 595 600 605  
 Val Lys Ala Glu Arg Asp Ile Leu Ala Glu Ala Asp Asn Glu Trp Val  
 610 615 620  
 Val Lys Leu Tyr Tyr Ser Phe Gln Asp Lys Asp Ser Leu Tyr Phe Val  
 625 630 635 640  
 Met Asp Tyr Ile Pro Gly Gly Asp Met Met Ser Leu Leu Ile Arg Met  
 645 650 655  
 Glu Val Phe Pro Glu His Leu Ala Arg Phe Tyr Ile Ala Glu Leu Thr  
 660 665 670  
 Leu Ala Ile Glu Ser Val His Lys Met Gly Phe Ile His Arg Asp Ile  
 675 680 685  
 Lys Pro Asp Asn Ile Leu Ile Asp Leu Asp Gly His Ile Lys Leu Thr  
 690 695 700  
 Asp Phe Gly Leu Cys Thr Gly Phe Arg Trp Thr His Asn Ser Lys Tyr  
 705 710 715 720  
 Tyr Gln Lys Gly Asn His Met Arg Gln Asp Ser Met Glu Pro Gly Asp  
 725 730 735

Leu Trp Asp Asp Val Ser Asn Cys Arg Cys Gly Asp Arg Leu Lys Thr  
                   740                                  745                                  750  
 Leu Glu Gln Arg Ala Gln Lys Gln His Gln Arg Cys Leu Ala His Ser  
                   755                                  760                                  765  
 Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu Val Leu Leu Arg Lys  
                   770                                  775                                  780  
 Gly Tyr Thr Gln Leu Cys Asp Trp Trp Ser Val Gly Val Ile Leu Phe  
                   785                                  790                                  795                                  800  
 Glu Met Leu Val Gly Gln Pro Pro Phe Leu Ala Pro Thr Pro Thr Glu  
                                   805                                  810                                  815  
 Thr Gln Leu Lys Val Ile Asn Trp Glu Ser Thr Leu His Ile Pro Thr  
                                   820                                  825                                  830  
 Gln Val Arg Leu Ser Ala Glu Ala Arg Asp Leu Ile Thr Lys Leu Cys  
                                   835                                  840                                  845  
 Cys Ala Ala Asp Cys Arg Leu Gly Arg Asp Gly Ala Asp Asp Leu Lys  
                                   850                                  855                                  860  
 Ala His Pro Phe Phe Asn Thr Ile Asp Phe Ser Arg Asp Ile Arg Lys  
                                   865                                  870                                  875                                  880  
 Gln Ala Ala Pro Tyr Val Pro Thr Ile Ser His Pro Met Asp Thr Ser  
                                   885                                  890                                  895  
 Asn Phe Asp Pro Val Asp Glu Glu Ser Pro Trp His Glu Ala Ser Gly  
                                   900                                  905                                  910  
 Glu Ser Ala Lys Ala Trp Asp Thr Leu Ala Ser Pro Ser Ser Lys His  
                                   915                                  920                                  925  
 Pro Glu His Ala Phe Tyr Glu Phe Thr Phe Arg Arg Phe Phe Asp Asp  
                                   930                                  935                                  940  
 Asn Gly Tyr Pro Phe Arg Cys Pro Lys Pro Ser Glu Pro Ala Glu Ser  
                                   945                                  950                                  955                                  960  
 Ala Asp Pro Gly Asp Ala Asp Leu Glu Gly Ala Ala Glu Gly Cys Gln  
                                   965                                  970                                  975  
 Pro Val Tyr Val  
                                   980

<210> 7  
 <211> 5720  
 <212> DNA  
 <213> Drosophila melanogaster

<220>  
 <221> CDS  
 <222> (1103)..(4399)

<400> 7  
atctagcacg acggcagcaa caaaaccacg aattaatttt actaaattta agccaaacgc 60  
gcatcggaata tgcccgaaaa tgccgattgaa tgcacgcgaa aagtgatggg ttgcgaacgc 120  
gagtgaatca agtgaaaata cgtcggcaaa tatcagcgaa ttgtcgtcaa aaggcaagga 180  
aaaacggaga aaaagaggaa aagcaataag tgccgtgtgt gggaaacgcg aaaaaggcga 240  
gaacaaagag gcgaaaagcg aggaaattgc gtggaaaacg tggaaaacgc gaagaagcga 300  
agctccaagt tgcccgccat cgattcgtgc gtaggatcaa ttaagattcc gagtggtcga 360  
gaatcggctc aaatcaaatt aaaatcaact aatatttttg tattcagata ttcaaattga 420  
attcattcat cgcccgcgac ttttattcgg atctgccaac tatttttgaa tttgaattgt 480  
gtgtctgcgg ctggcgcgaga atctctgata aagcagagga ataaaatcgg aagaacaaca 540  
aatacaataa caaatgaaat gcgggggagca gtatttacat gccaaatgaa tgctggatag 600  
gcgaaagggg gggtttctct tataatgcaa atgtgaatgt gaatgcgaat gcgaatgcga 660  
gtggaagaat tcccggcgcg agtgataaat aatccgacga caaacaagc agaagcctac 720  
accgcgagaa agagcagcgc aaacacaatt atctttattg agagcaacaa tatcaagatc 780  
gagataataa agcatcctaa aaccgcgcgc ttagttcgtt ttagtctcgc cacggatata 840  
gatattcaaa ggcaaaaagg tgggtgtcggc atcgccagac aaacaagtaa agcatctatt 900  
tcatacaaaa caaccaatta aataataata aaaataataa taatcgtaga gaggcagagc 960  
caaatcaaat tcccggcgcg cgatgtgccc cagtgtgtgt gcgtgtgtgt gtgtgtgtgc 1020  
tgtgtgtgtc tgtgcgagtg ttagtgtgcg gagcatttct gtgatatgag tgctaaatgc 1080  
cacagggcga agcagcagca tc atg cat cca gcg ggc gaa aaa agg ggc ggt 1132  
Met His Pro Ala Gly Glu Lys Arg Gly Gly  
1 5 10  
cgc ccc aat gat aaa tac acg gcg gaa gcc ctc gag agc atc aag cag 1180  
Arg Pro Asn Asp Lys Tyr Thr Ala Glu Ala Leu Glu Ser Ile Lys Gln  
15 20 25  
gac cta acc cga ttt gaa gta caa aat aac cat agg aat aat cag aat 1228  
Asp Leu Thr Arg Phe Glu Val Gln Asn Asn His Arg Asn Asn Gln Asn  
30 35 40  
tac aca cct ctg cga tac acg gcg acc aac gga cgc aac gat gca ctt 1276  
Tyr Thr Pro Leu Arg Tyr Thr Ala Thr Asn Gly Arg Asn Asp Ala Leu  
45 50 55  
act cct gac tat cac cac gcc aag cag ccg atg gag ccg cca ccc tcc 1324  
Thr Pro Asp Tyr His His Ala Lys Gln Pro Met Glu Pro Pro Pro Ser  
60 65 70

gcc tct cct gct ccg gac gtg gtc ata ccg ccg ccg gcc att gta	1372
Ala Ser Pro Ala Pro Asp Val Val Ile Pro Pro Pro Ala Ile Val	
75 80 85 90	
ggt cag ccc gga gcc ggc tcc ata tcc gta tcc ggt gtg ggc gtt gga	1420
Gly Gln Pro Gly Ala Gly Ser Ile Ser Val Ser Gly Val Gly Val Gly	
95 100 105	
gtg gtg ggt gtg gcg aac gga cgt gtg cca aag atg atg acg gcc cta	1468
Val Val Gly Val Ala Asn Gly Arg Val Pro Lys Met Met Thr Ala Leu	
110 115 120	
atg cca aac aaa ctg atc cgg aag ccg agc atc gaa cgg gac acg gcg	1516
Met Pro Asn Lys Leu Ile Arg Lys Pro Ser Ile Glu Arg Asp Thr Ala	
125 130 135	
agc agt cac tac ctg cgc tgc agt ccg gct ctg gac tcc gga gcc ggt	1564
Ser Ser His Tyr Leu Arg Cys Ser Pro Ala Leu Asp Ser Gly Ala Gly	
140 145 150	
agc tcc cga tcg gac agc ccc cat tcg cac cac acc cac cag ccg agc	1612
Ser Ser Arg Ser Asp Ser Pro His Ser His His Thr His Gln Pro Ser	
155 160 165 170	
tcg agg acg gtg ggt aat cca ggt gga aat ggt gga ttt tct ccg tcg	1660
Ser Arg Thr Val Gly Asn Pro Gly Gly Asn Gly Gly Phe Ser Pro Ser	
175 180 185	
cca agc ggt ttc agt gag gtg gct cca ccg gcg ccg ccg cca cgc aat	1708
Pro Ser Gly Phe Ser Glu Val Ala Pro Pro Ala Pro Pro Pro Arg Asn	
190 195 200	
ccc acc gcc tcc agc gcg gcc acg ccc cca ccg cca gtg ccg ccc acc	1756
Pro Thr Ala Ser Ser Ala Ala Thr Pro Pro Pro Pro Val Pro Pro Thr	
205 210 215	
agc cag gcg tac gtg aag cgg cga tca ccg gcc ctg aac aac cgc ccg	1804
Ser Gln Ala Tyr Val Lys Arg Arg Ser Pro Ala Leu Asn Asn Arg Pro	
220 225 230	
ccg gcg ata gcg cca ccc act cag cga ggc aac tca cct gta ata acc	1852
Pro Ala Ile Ala Pro Pro Thr Gln Arg Gly Asn Ser Pro Val Ile Thr	
235 240 245 250	
caa aac ggg ctg aag aac ccg cag cag cag ttg acg cag cag ctg aag	1900
Gln Asn Gly Leu Lys Asn Pro Gln Gln Gln Leu Thr Gln Gln Leu Lys	
255 260 265	
tcc ctg aac cta tac cca ggc gga ggc agt gga gca gtg gtg gag cca	1948
Ser Leu Asn Leu Tyr Pro Gly Gly Gly Ser Gly Ala Val Val Glu Pro	
270 275 280	
ccg ccg ccc tac cta att caa ggc gga gcc gga gga gca gca ccg ccg	1996
Pro Pro Pro Tyr Leu Ile Gln Gly Gly Ala Gly Gly Ala Ala Pro Pro	
285 290 295	

ccg cca cca ccc agt tac acg gcc tcc atg cag tcg cgg cag tcg ccc Pro Pro Pro Pro Ser Tyr Thr Ala Ser Met Gln Ser Arg Gln Ser Pro 300 305 310	2044
aca caa tcc caa caa tcg gac tac agg aaa tcc ccg agc agt ggg ata Thr Gln Ser Gln Gln Ser Asp Tyr Arg Lys Ser Pro Ser Ser Gly Ile 315 320 325 330	2092
tac tcg gcc acc tcg gcg ggc tcg ccg agc ccc ata act gtg tcg ctg Tyr Ser Ala Thr Ser Ala Gly Ser Pro Ser Pro Ile Thr Val Ser Leu 335 340 345	2140
ccg ccg gcg ccg ctg gcg aag cca caa cca cga gtc tac cag gcc agg Pro Pro Ala Pro Leu Ala Lys Pro Gln Pro Arg Val Tyr Gln Ala Arg 350 355 360	2188
agt cag cag ccg atc atc atg cag agt gtg aag agc acg cag gtc caa Ser Gln Gln Pro Ile Ile Met Gln Ser Val Lys Ser Thr Gln Val Gln 365 370 375	2236
aag ccc gtg ctg caa aca gca gtg gcg cgc caa tcg cca tcg agt gcc Lys Pro Val Leu Gln Thr Ala Val Ala Arg Gln Ser Pro Ser Ser Ala 380 385 390	2284
tcg gcc agc aat tca cca gtc cac gtg ctg gcc gct cca ccc tct tac Ser Ala Ser Asn Ser Pro Val His Val Leu Ala Ala Pro Pro Ser Tyr 395 400 405 410	2332
cct cag aag tcc gcg gca gtg gtg cag cag cag caa cag gca gca gcg Pro Gln Lys Ser Ala Ala Val Val Gln Gln Gln Gln Gln Ala Ala Ala 415 420 425	2380
gcg gcc cac cag cag cag cat cag cac cag caa tcc aaa cca cca acg Ala Ala His Gln Gln Gln His Gln Gln Gln Ser Lys Pro Pro Thr 430 435 440	2428
cca acc aca ccg ccc ttg gtg ggt ctg aac agc aag ccc aat tgc ctg Pro Thr Thr Pro Pro Leu Val Gly Leu Asn Ser Lys Pro Asn Cys Leu 445 450 455	2476
gag cca ccg tcc tat gcc aag agc atg cag gcc aag gcg gcc acg gtg Glu Pro Pro Ser Tyr Ala Lys Ser Met Gln Ala Lys Ala Ala Thr Val 460 465 470	2524
gta cag cag cag caa cag cag cag caa caa cag cag gtc cag cag cag Val Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Val Gln Gln Gln 475 480 485 490	2572
cag gtg caa cag cag cag caa cag cag caa cag caa ctg cag gcc ttg Gln Val Gln Gln Gln Gln Gln Gln Gln Gln Gln Leu Gln Ala Leu 495 500 505	2620
agg gtg ctc cag gca cag gct cag agg gag cgg gat caa cgg gag cgg Arg Val Leu Gln Ala Gln Ala Gln Arg Glu Arg Asp Gln Arg Glu Arg 510 515 520	2668

gaa cgg gat cag cag aag ctg gcc aac gga aat cct ggc cgg cag atg Glu Arg Asp Gln Gln Lys Leu Ala Asn Gly Asn Pro Gly Arg Gln Met 525 530 535	2716
ctt ccg ccg ccg ccc tat cag agc aac aac aac aac aac agc gag atc Leu Pro Pro Pro Pro Tyr Gln Ser Asn Asn Asn Asn Asn Ser Glu Ile 540 545 550	2764
aaa ccg ccg agc tgc aac aac aac aac ata cag ata agc aac agc aac Lys Pro Pro Ser Cys Asn Asn Asn Asn Ile Gln Ile Ser Asn Ser Asn 555 560 565 570	2812
ctg gcg acg aca cca ccc att ccg cct gcc aaa tac aat aac aac tcc Leu Ala Thr Thr Pro Pro Ile Pro Pro Ala Lys Tyr Asn Asn Asn Ser 575 580 585	2860
tcc aac acg ggc gcg aat agc tcg ggc ggc agc aac gga tcc acc ggc Ser Asn Thr Gly Ala Asn Ser Ser Gly Ser Asn Gly Ser Thr Gly 590 595 600	2908
acc acc gcc tcc tcg tcg acc agc tgc aag aag atc aag cac gcc tcg Thr Thr Ala Ser Ser Ser Thr Ser Cys Lys Lys Ile Lys His Ala Ser 605 610 615	2956
ccc atc ccg gag cgc aag aag atc tcc aag gag aag gag gag gag cgc Pro Ile Pro Glu Arg Lys Lys Ile Ser Lys Glu Lys Glu Glu Glu Arg 620 625 630	3004
aag gag ttc cgc atc agg cag tac tcg ccg caa gcc ttc aag ttc ttc Lys Glu Phe Arg Ile Arg Gln Tyr Ser Pro Gln Ala Phe Lys Phe Phe 635 640 645 650	3052
atg gag cag cac ata gag aac gtg atc aag tcg tat cgc cag cgc acg Met Glu Gln His Ile Glu Asn Val Ile Lys Ser Tyr Arg Gln Arg Thr 655 660 665	3100
tat cgc aag aat cag ctg gag aag gag atg cac aaa gtg gga ctg ccc Tyr Arg Lys Asn Gln Leu Glu Lys Glu Met His Lys Val Gly Leu Pro 670 675 680	3148
gat cag acc caa atc gag atg agg aaa atg ctg aac caa aag gag agc Asp Gln Thr Gln Ile Glu Met Arg Lys Met Leu Asn Gln Lys Glu Ser 685 690 695	3196
aac tac att cga ttg aag cgc gcc aag atg gac aag agc atg ttc gtc Asn Tyr Ile Arg Leu Lys Arg Ala Lys Met Asp Lys Ser Met Phe Val 700 705 710	3244
aaa ctg aag ccc att gga gtg ggt gca ttt ggc gag gta acg ctg gtg Lys Leu Lys Pro Ile Gly Val Gly Ala Phe Gly Glu Val Thr Leu Val 715 720 725 730	3292
agc aaa atc gat acc tcg aac cat ttg tat gcg atg aaa acc ctg cgg Ser Lys Ile Asp Thr Ser Asn His Leu Tyr Ala Met Lys Thr Leu Arg 735 740 745	3340



aaa gcg gac gtt ctc aag cgg aat cag gtg gca cac gtg aag gcc gag	3388
Lys Ala Asp Val Leu Lys Arg Asn Gln Val Ala His Val Lys Ala Glu	
750 755 760	
agg gat atc ctc gcg gaa gcc gac aat aac tgg gtg gtg aag ttg tac	3436
Arg Asp Ile Leu Ala Glu Ala Asp Asn Asn Trp Val Val Lys Leu Tyr	
765 770 775	
tac agc ttc cag gac aag gat aat ctg tac ttt gtg atg gac tac ata	3484
Tyr Ser Phe Gln Asp Lys Asp Asn Leu Tyr Phe Val Met Asp Tyr Ile	
780 785 790	
cca ggt ggt gat ctg atg tcg ctg ctc atc aaa ctg ggc att ttc gag	3532
Pro Gly Gly Asp Leu Met Ser Leu Leu Ile Lys Leu Gly Ile Phe Glu	
795 800 805 810	
gag gaa ctg gcc aga ttc tac atc gcc gag gtc acc tgc gcc gtg gac	3580
Glu Glu Leu Ala Arg Phe Tyr Ile Ala Glu Val Thr Cys Ala Val Asp	
815 820 825	
agc gtt cac aaa atg ggc ttc att cac aga gac atc aag cct gac aac	3628
Ser Val His Lys Met Gly Phe Ile His Arg Asp Ile Lys Pro Asp Asn	
830 835 840	
ata ctc atc gat agg gac gga cac ata aag ctg acc gac ttt ggc ctg	3676
Ile Leu Ile Asp Arg Asp Gly His Ile Lys Leu Thr Asp Phe Gly Leu	
845 850 855	
tgc acg gga ttc cga tgg acg cac aac tcg aag tac tac cag gag aac	3724
Cys Thr Gly Phe Arg Trp Thr His Asn Ser Lys Tyr Tyr Gln Glu Asn	
860 865 870	
ggc aat cac tcg cgc cag gac tcg atg gag ccc tgg gag gaa tac tcc	3772
Gly Asn His Ser Arg Gln Asp Ser Met Glu Pro Trp Glu Glu Tyr Ser	
875 880 885 890	
gag aac gga ccg aag ccc acc gtg ctg gag agg cga cgg atg cgc gat	3820
Glu Asn Gly Pro Lys Pro Thr Val Leu Glu Arg Arg Arg Met Arg Asp	
895 900 905	
cac caa aga gtc ctg gcc cac tcg ctg gtg ggc acc ccg aac tac ata	3868
His Gln Arg Val Leu Ala His Ser Leu Val Gly Thr Pro Asn Tyr Ile	
910 915 920	
gct ccc gag gtg ctg gag agg agt ggg tac acg cag ctg tgc gac tac	3916
Ala Pro Glu Val Leu Glu Arg Ser Gly Tyr Thr Gln Leu Cys Asp Tyr	
925 930 935	
tgg agc gtg ggc gtc atc ctt tac gag atg ctg gtg ggt cag ccg ccc	3964
Trp Ser Val Gly Val Ile Leu Tyr Glu Met Leu Val Gly Gln Pro Pro	
940 945 950	
ttt ctg gcc aac agt ccg ctg gaa acg caa caa aag gtc atc aac tgg	4012
Phe Leu Ala Asn Ser Pro Leu Glu Thr Gln Gln Lys Val Ile Asn Trp	
955 960 965 970	

gag aaa acg ctg cat att ccg ccg cag gcc gag tta tcc cgc gag gct 4060  
 Glu Lys Thr Leu His Ile Pro Pro Gln Ala Glu Leu Ser Arg Glu Ala  
 975 980 985

acg gac ttg ata agg agg ctc tgt gcg tcg gct gac aag cgg ctg ggc 4108  
 Thr Asp Leu Ile Arg Arg Leu Cys Ala Ser Ala Asp Lys Arg Leu Gly  
 990 995 1000

aag agc gtg gac gag gtc aag agc cac gac ttc ttc aag ggc atc gac 4156  
 Lys Ser Val Asp Glu Val Lys Ser His Asp Phe Phe Lys Gly Ile Asp  
 1005 1010 1015

ttt gcg gac atg cgg aag cag aaa gcg ccc tac ata ccg gaa atc aag 4204  
 Phe Ala Asp Met Arg Lys Gln Lys Ala Pro Tyr Ile Pro Glu Ile Lys  
 1020 1025 1030

cac cca acg gac aca tcc aac ttt gat ccc gtg gat ccg gag aag ctg 4252  
 His Pro Thr Asp Thr Ser Asn Phe Asp Pro Val Asp Pro Glu Lys Leu  
 1035 1040 1045 1050

cgc tcg aat gac tcc acc atg agc agc ggc gat gat gtc gac cag aat 4300  
 Arg Ser Asn Asp Ser Thr Met Ser Ser Gly Asp Asp Val Asp Gln Asn  
 1055 1060 1065

gac cgc act ttc cac ggc ttt ttc gaa ttt acc ttc cgt cgc ttc ttc 4348  
 Asp Arg Thr Phe His Gly Phe Phe Glu Phe Thr Phe Arg Arg Phe Phe  
 1070 1075 1080

gac gac aag cag ccg ccg gat atg acg gac gat cag gcg ccg gtt tac 4396  
 Asp Asp Lys Gln Pro Pro Asp Met Thr Asp Asp Gln Ala Pro Val Tyr  
 1085 1090 1095

gtc tgaaatggat gctctccatg tgcccaacac caacaccccg ccccccgaatc 4449  
 Val

attgttagtc aaatagtcac aaaaagggga tagaaaccat tgagtgggct tgcattgtaa 4509

aggaagcgtg gctatagaat gaaactatct atatacatta tataaattat aggagacagt 4569

agaggcgagg gctacgtata tacatacaaa taatatacat atatttgata tatatatata 4629

tatatatgcc gtagggcatg aactgaataa atataaaacg gagccgagta gagatgaaac 4689

gagaggagcg agtcaggacc ttcgaccttt aactgaacat agtatatcct tgtgcactac 4749

tactccacaa caaatatata tttttaaatt gttagaattc aaaaggggacc aactggaaat 4809

cgaacctttc tggtgctcaa agcaaagcaa agcaaagcaa aacaaaacgc cttaaactaa 4869

atgagacgcg aatttaccca accacttcac tcctctcctt tctccacctc cgatcggtgg 4929

ccggattcga actcagcagg ctggttgcat ccggccatcc cattgacttc ccattcagaa 4989

ttgagattgc gaggtgtgcg atggagaacg aacggagacc aaaagtcgca cggcagcgat 5049

ataagcgggt cttataagcc taatctaaat ctaaactggg agaacaggac ctatgtatgt 5109

cctgctatcc aattcgtcta tcaactgctct tcactgtgtg acgaccccca cccccccct 5169



```

ccccatccaa aagaacaaac ttagacgtag cctatgtgaa aagctagcaa tgtagacca 5229
acttggtgaa tgccaaatga aattgttttag cccacgagg aaaacgcggg ggaaattcaa 5289
cacttattct ctgatagcaa acggaaaaga aagaaagaaa aaaaaaaca gaaacagtac 5349
gagaaaattg taatcttctt aatgtaatat tgtaaagaac acgttaattg taatctatgc 5409
tagagttgtg tagcgcccta agatgttttt tagtttatag accgctaacc gtaatctagt 5469
ttaattccta acactaagcg agagtacagt acattgggtt ttttgttgt cgtaggttcg 5529
ttggaaaatg cttaacggga aacgatttgt ttttctctt aattagcttc agtttgtatg 5589
tgcgtgtgtt tttattatga cttatatata gtccatctga atattcgtgg atggagccta 5649
ttttaaatgt gagatcgagc taattgaagg aaatacaaac aaactctgtg tgccttggcc 5709
aattagttta c

```

5720

```

<210> 8
<211> 1099
<212> PRT
<213> Drosophila melanogaster

```

```

<400> 8
Met His Pro Ala Gly Glu Lys Arg Gly Gly Arg Pro Asn Asp Lys Tyr
  1             5             10             15
Thr Ala Glu Ala Leu Glu Ser Ile Lys Gln Asp Leu Thr Arg Phe Glu
          20             25             30
Val Gln Asn Asn His Arg Asn Asn Gln Asn Tyr Thr Pro Leu Arg Tyr
      35             40             45
Thr Ala Thr Asn Gly Arg Asn Asp Ala Leu Thr Pro Asp Tyr His His
      50             55             60
Ala Lys Gln Pro Met Glu Pro Pro Pro Ser Ala Ser Pro Ala Pro Asp
      65             70             75             80
Val Val Ile Pro Pro Pro Pro Ala Ile Val Gly Gln Pro Gly Ala Gly
          85             90             95
Ser Ile Ser Val Ser Gly Val Gly Val Gly Val Val Gly Val Ala Asn
      100             105             110
Gly Arg Val Pro Lys Met Met Thr Ala Leu Met Pro Asn Lys Leu Ile
      115             120             125
Arg Lys Pro Ser Ile Glu Arg Asp Thr Ala Ser Ser His Tyr Leu Arg
      130             135             140
Cys Ser Pro Ala Leu Asp Ser Gly Ala Gly Ser Ser Arg Ser Asp Ser
      145             150             155             160

```

Pro His Ser His His Thr His Gln Pro Ser Ser Arg Thr Val Gly Asn  
 165 170 175  
 Pro Gly Gly Asn Gly Gly Phe Ser Pro Ser Pro Ser Gly Phe Ser Glu  
 180 185 190  
 Val Ala Pro Pro Ala Pro Pro Pro Arg Asn Pro Thr Ala Ser Ser Ala  
 195 200 205  
 Ala Thr Pro Pro Pro Pro Val Pro Pro Thr Ser Gln Ala Tyr Val Lys  
 210 215 220  
 Arg Arg Ser Pro Ala Leu Asn Asn Arg Pro Pro Ala Ile Ala Pro Pro  
 225 230 235 240  
 Thr Gln Arg Gly Asn Ser Pro Val Ile Thr Gln Asn Gly Leu Lys Asn  
 245 250 255  
 Pro Gln Gln Gln Leu Thr Gln Gln Leu Lys Ser Leu Asn Leu Tyr Pro  
 260 265 270  
 Gly Gly Gly Ser Gly Ala Val Val Glu Pro Pro Pro Pro Tyr Leu Ile  
 275 280 285  
 Gln Gly Gly Ala Gly Gly Ala Ala Pro Pro Pro Pro Pro Pro Ser Tyr  
 290 295 300  
 Thr Ala Ser Met Gln Ser Arg Gln Ser Pro Thr Gln Ser Gln Gln Ser  
 305 310 315 320  
 Asp Tyr Arg Lys Ser Pro Ser Ser Gly Ile Tyr Ser Ala Thr Ser Ala  
 325 330 335  
 Gly Ser Pro Ser Pro Ile Thr Val Ser Leu Pro Pro Ala Pro Leu Ala  
 340 345 350  
 Lys Pro Gln Pro Arg Val Tyr Gln Ala Arg Ser Gln Gln Pro Ile Ile  
 355 360 365  
 Met Gln Ser Val Lys Ser Thr Gln Val Gln Lys Pro Val Leu Gln Thr  
 370 375 380  
 Ala Val Ala Arg Gln Ser Pro Ser Ser Ala Ser Ala Ser Asn Ser Pro  
 385 390 395 400  
 Val His Val Leu Ala Ala Pro Pro Ser Tyr Pro Gln Lys Ser Ala Ala  
 405 410 415  
 Val Val Gln Gln Gln Gln Gln Ala Ala Ala Ala His Gln Gln Gln  
 420 425 430  
 His Gln His Gln Gln Ser Lys Pro Pro Thr Pro Thr Thr Pro Pro Leu  
 435 440 445  
 Val Gly Leu Asn Ser Lys Pro Asn Cys Leu Glu Pro Pro Ser Tyr Ala  
 450 455 460

Lys Ser Met Gln Ala Lys Ala Ala Thr Val Val Gln Gln Gln Gln Gln  
 465 470 475 480  
 Gln Gln Gln Gln Gln Gln Val Gln Gln Gln Gln Val Gln Gln Gln Gln  
 485 490 495  
 Gln Gln Gln Gln Gln Gln Leu Gln Ala Leu Arg Val Leu Gln Ala Gln  
 500 505 510  
 Ala Gln Arg Glu Arg Asp Gln Arg Glu Arg Glu Arg Asp Gln Gln Lys  
 515 520 525  
 Leu Ala Asn Gly Asn Pro Gly Arg Gln Met Leu Pro Pro Pro Pro Tyr  
 530 535 540  
 Gln Ser Asn Asn Asn Asn Asn Ser Glu Ile Lys Pro Pro Ser Cys Asn  
 545 550 555 560  
 Asn Asn Asn Ile Gln Ile Ser Asn Ser Asn Leu Ala Thr Thr Pro Pro  
 565 570 575  
 Ile Pro Pro Ala Lys Tyr Asn Asn Asn Ser Ser Asn Thr Gly Ala Asn  
 580 585 590  
 Ser Ser Gly Gly Ser Asn Gly Ser Thr Gly Thr Thr Ala Ser Ser Ser  
 595 600 605  
 Thr Ser Cys Lys Lys Ile Lys His Ala Ser Pro Ile Pro Glu Arg Lys  
 610 615 620  
 Lys Ile Ser Lys Glu Lys Glu Glu Glu Arg Lys Glu Phe Arg Ile Arg  
 625 630 635 640  
 Gln Tyr Ser Pro Gln Ala Phe Lys Phe Phe Met Glu Gln His Ile Glu  
 645 650 655  
 Asn Val Ile Lys Ser Tyr Arg Gln Arg Thr Tyr Arg Lys Asn Gln Leu  
 660 665 670  
 Glu Lys Glu Met His Lys Val Gly Leu Pro Asp Gln Thr Gln Ile Glu  
 675 680 685  
 Met Arg Lys Met Leu Asn Gln Lys Glu Ser Asn Tyr Ile Arg Leu Lys  
 690 695 700  
 Arg Ala Lys Met Asp Lys Ser Met Phe Val Lys Leu Lys Pro Ile Gly  
 705 710 715 720  
 Val Gly Ala Phe Gly Glu Val Thr Leu Val Ser Lys Ile Asp Thr Ser  
 725 730 735  
 Asn His Leu Tyr Ala Met Lys Thr Leu Arg Lys Ala Asp Val Leu Lys  
 740 745 750  
 Arg Asn Gln Val Ala His Val Lys Ala Glu Arg Asp Ile Leu Ala Glu  
 755 760 765

Ala Asp Asn Asn Trp Val Val Lys Leu Tyr Tyr Ser Phe Gln Asp Lys  
 770 775 780  
 Asp Asn Leu Tyr Phe Val Met Asp Tyr Ile Pro Gly Gly Asp Leu Met  
 785 790 795 800  
 Ser Leu Leu Ile Lys Leu Gly Ile Phe Glu Glu Glu Leu Ala Arg Phe  
 805 810 815  
 Tyr Ile Ala Glu Val Thr Cys Ala Val Asp Ser Val His Lys Met Gly  
 820 825 830  
 Phe Ile His Arg Asp Ile Lys Pro Asp Asn Ile Leu Ile Asp Arg Asp  
 835 840 845  
 Gly His Ile Lys Leu Thr Asp Phe Gly Leu Cys Thr Gly Phe Arg Trp  
 850 855 860  
 Thr His Asn Ser Lys Tyr Tyr Gln Glu Asn Gly Asn His Ser Arg Gln  
 865 870 875 880  
 Asp Ser Met Glu Pro Trp Glu Glu Tyr Ser Glu Asn Gly Pro Lys Pro  
 885 890 895  
 Thr Val Leu Glu Arg Arg Arg Met Arg Asp His Gln Arg Val Leu Ala  
 900 905 910  
 His Ser Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu Val Leu Glu  
 915 920 925  
 Arg Ser Gly Tyr Thr Gln Leu Cys Asp Tyr Trp Ser Val Gly Val Ile  
 930 935 940  
 Leu Tyr Glu Met Leu Val Gly Gln Pro Pro Phe Leu Ala Asn Ser Pro  
 945 950 955 960  
 Leu Glu Thr Gln Gln Lys Val Ile Asn Trp Glu Lys Thr Leu His Ile  
 965 970 975  
 Pro Pro Gln Ala Glu Leu Ser Arg Glu Ala Thr Asp Leu Ile Arg Arg  
 980 985 990  
 Leu Cys Ala Ser Ala Asp Lys Arg Leu Gly Lys Ser Val Asp Glu Val  
 995 1000 1005  
 Lys Ser His Asp Phe Phe Lys Gly Ile Asp Phe Ala Asp Met Arg Lys  
 1010 1015 1020  
 Gln Lys Ala Pro Tyr Ile Pro Glu Ile Lys His Pro Thr Asp Thr Ser  
 1025 1030 1035 1040  
 Asn Phe Asp Pro Val Asp Pro Glu Lys Leu Arg Ser Asn Asp Ser Thr  
 1045 1050 1055  
 Met Ser Ser Gly Asp Asp Val Asp Gln Asn Asp Arg Thr Phe His Gly  
 1060 1065 1070

Phe Phe Glu Phe Thr Phe Arg Arg Phe Phe Asp Asp Lys Gln Pro Pro  
 1075 1080 1085

Asp Met Thr Asp Asp Gln Ala Pro Val Tyr Val  
 1090 1095

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/19068

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(C) : Please See Extra Sheet.

US CL : 424/130.1; 536/23.5; 514/1, 2, 44; 800/3, 13, 14, 21, 25

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/130.1; 536/23.5; 514/1, 2, 44; 800/3, 13, 14, 21, 25

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, STN, MEDLINE, CAPLUS, BIOSIS, SCISEARCH

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96/30402 A1 (YALE UNIVERSITY) 03 October 1996, entire document, especially pages 158-163.	1-5
Y		6-113
Y	WO 95/31722 A1 (LIGAND PHARMACEUTICALS, INC.) 23 November 1995, entire document, especially page 50.	6-37
Y, P	TAO et al. Human homologue of the Drosophila melanogaster lats tumor suppressor modulates CDC2 activity, Nature Genetics, February 1999, Vol 21, No. 2, page 177-181, entire document.	68-113

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*g* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

27 OCTOBER 1999

Date of mailing of the international search report

16 DEC 1999

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

SHIN-LIN CHEN

Telephone No. (703) 308-0196



**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/US99/19068

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	ST. JOHN et al. Mice Deficient of Lats1 Develop Soft-Tissue Sarcomas, Ovarian Tumors and Pituitary Dysfunction, Nature Genetics, February 1999, Vol 21, No. 2, page 182-186, entire document.	1-37

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/19068

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/19068

## A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

A61K 39/395; C07H 21/04; A01N 61/00, 37/18, 43/04; C12N 5/00, 15/00



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification:</b> <b>G01N 33/50, 33/68</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 95/31722</b> <b>(43) International Publication Date:</b> 23 November 1995 (23.11.95)
<b>(21) International Application Number:</b> PCT/US95/06524 <b>(22) International Filing Date:</b> 15 May 1995 (15.05.95)  <b>(30) Priority Data:</b> 08/245,470 18 May 1994 (18.05.94) US  <b>(71) Applicant:</b> LIGAND PHARMACEUTICALS, INC. [US/US]; 9393 Towne Centre Drive, San Diego, CA 92121 (US).  <b>(72) Inventors:</b> HERMANN, Thomas; 950 Santa Helena Park Court, Solana Beach, CA 92075 (US). PIKE, John, W.; 911 Springwood Lane, Encinitas, CA 92024 (US).  <b>(74) Agents:</b> WARBURG, Richard, J. et al.; Lyon & Lyon, First Interstate World Center, Suite 4700, 633 West Fifth Street, Los Angeles, CA 90071-2066 (US).		<b>(81) Designated States:</b> AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MW, MX, NO, NZ, PL, RO, RU, SD, SG, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> SCREENING FOR CYTOKINE MODULATORS  <b>(57) Abstract</b>  This invention provides a method for screening for agents useful for treatment of diseases and pathological conditions affected by cytokines. These agents interact directly or indirectly with an intracellular receptor, which in turn modulates the binding of a rel-like protein, a rel-like protein complex, or other transcriptional proteins to a rel site on the promoter of a cytokine gene. The intracellular receptor can be the estrogen receptor, retinoid acid receptors, retinoid X receptors, glucocorticoid receptor, progesterone receptors, androgen receptor, thyroid hormone receptors, or vitamin D receptor. The select agents can be used to treat osteoporosis, rheumatoid arthritis, inflammation, psoriasis, Kaposi's sarcoma, septic shock and multiple myeloma.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## SCREENING FOR CYTOKINE MODULATORS

FIELD OF THE INVENTION

5 This invention relates to a method for screening for agents useful for treatment of diseases and pathological conditions affected by cytokines and novel agents identified using such screening method.

BACKGROUND OF THE INVENTION

10 Cytokines are a group of molecules capable of signalling cellular development. Aberrant expression of cytokines is known to be associated with pathological conditions including autoimmune diseases, septic shock, rheumatoid arthritis, psoriasis, inflammation,  
15 postmenopausal osteoporosis, and some cancers. Common treatment for these pathological conditions are retinoids, immunosuppressants, glucocorticoids and other steroid drugs. Estrogens are specifically employed in the prevention of postmenopausal osteoporosis.

20 Steroids and related hormone drugs exert their therapeutic effects by binding to a superfamily of intracellular receptors (IRs), which are regulators of gene transcription. IRs can function as activators as well as repressors of specific cytokine genes. The  
25 activity of IRs is controlled by hormones or other ligands that bind to the IRs.

The classical mechanism of transcriptional regulation by IRs involves binding of the IRs to specific response elements in the promoters of the  
30 regulated genes, for example, the binding of the estrogen receptor to its response site in the vitellogenin gene (Klein-Hitpass et al., Cell 46:1053-1061, 1986). More recently a different mechanism of IRs function has been described in glucocorticoid receptor  
35 mediated AP-1 transcription regulation that does not

require direct DNA-binding of the IRs (Yang-Yen et al., Cell 62:1205-1215, 1990).

Although steroid drugs have been shown to repress the level of certain cytokines, a lack of tissue specificity and side effects of the steroids may limit their use as therapeutic agents. These side effects may be reduced or completely avoided with more specifically acting compounds.

Pfahl and Karin (PCT publication, WO 92/07072, 1992) describes a method of screening a sample for ligands which bind to a nuclear receptor to form a complex which binds or interferes with a non rel-like protein AP-1 or an AP-1 component.

#### SUMMARY OF THE INVENTION

The present invention relates to a method for identifying new therapeutic agents and for using these agents to treat diseases and conditions affected by cytokines, such as, but not limited to, osteoporosis, rheumatoid arthritis, inflammation, psoriasis, septic shock, Kaposi's sarcoma and multiple myeloma. This method makes it possible to screen large collections of natural, semisynthetic, or synthetic compounds for therapeutic agents that affect the transcription of a cytokine through an intracellular receptor mediated pathway.

By "cytokine" is generally meant a secreted protein which acts as a chemical mediator of cellular regulation. More specifically, it is meant a diverse groups of soluble polypeptides such as growth factors and hormones that control the growth, differentiation and function of cells, including, but not limited to, GM-CSF, G-CSF, IL-2, IL-6, IL-8, and IL-11.

The present invention relates to the determination that inhibition of interleukin 6 (IL-6) expression by estrogen-estrogen receptor complex is mediated through

the control of the transcriptional activity of NF $\kappa$ B or closely related proteins on the IL-6 promoter. This mechanism does not involve direct binding of ER to IL-6 promoter but controls the DNA-binding properties of the activated NF $\kappa$ B and possible other members of the rel-family of proteins to their specific response elements (i.e., rel site) on the IL-6 promoter.

Because NF $\kappa$ B is involved in the regulation of genes encoding various cytokines and their receptors, viral proteins, and proteins involved in the acute-phase response, the regulation of NF $\kappa$ B activity by estrogen and possible other hormones is of general importance (see generally Baeuerle, *Biochemica et Biophysica Acta*, 1072:63-80, 1993, incorporated by reference herein). For example, retinoic acid treatment, which strongly inhibits IL-6 expression in +/+LDA11 cells and other tissues (Gross, V., P. M. Villiger, B. Zhang, and M. Lotz, 1993, "Retinoic acid inhibits interleukin-1-induced cytokine synthesis in human monocytes," J. Leukoc. Biol. 54:125-132), has the same effect as estrogen on the NF $\kappa$ B related complexes with the IL-6 promoter. This suggests a general pathway of transcriptional regulation involving cross-talk between members of the intracellular receptor family and the NF $\kappa$ B transcription factors.

The above determination allows for the screening of drugs that specifically influence genes controlled by the rel-transcription factors, i.e. genes involved in inflammation, sepsis, skin and kidney disorders, osteoporosis, certain cancers, and hematopoietic dysfunctions without the side effects of known steroid drugs. The diseases listed are usually correlated with aberrant expression of cytokines such as IL-1, TNF $\alpha$ , IL-6, IL-8 that are under the control of NF $\kappa$ B or other rel proteins.



Thus, the present invention features a method for identifying agents which, by activating an intracellular receptor, cause a significant reduction in the binding of a rel-like protein or other transcriptional protein to the rel site on the promoter of a cytokine gene or a portion of the promoter, thereby reducing the transcription of the cytokine.

By "intracellular receptor" is meant an intracellular transcription factor whose activity is regulated by binding of small molecules, including, but not limited to, estrogen receptor, retinoid acid receptors, retinoid X receptors, glucocorticoid receptor, progesterone receptors, androgen receptor, thyroid hormone receptors, and vitamin D receptor.

By "rel-like protein" is meant a protein or a protein complex of the rel family that share a homology in the rel domain and is involved in gene regulation (see Liou and Baltimore, Current Opinion in Cell Biology, 5:477-487, 1993, incorporated by reference herein), including, but not limited to, NF $\kappa$ B, L $\gamma$ t-10, c-rel, and relB.

By "transcriptional protein" is meant a cytoplasmic or nuclear protein that, when activated, bind a promoter either directly, or indirectly through a complex of proteins to modulate the transcription activity of the promoter.

By "rel site" is meant a DNA sequence that serves as a binding site for rel-like proteins or complexes comprising one or more rel-like proteins, including, but not limited to,  $\kappa$ B motifs identified in Baeuerle, Biochemica et Biophysica Acta, 1072:63-80, 1993, incorporated by reference herein, such as the NF $\kappa$ B binding site on IL-6 promoter.

By "promoter" is meant a DNA regulatory region capable of binding directly or indirectly to RNA polymerase in a cell and initiating transcription of a

downstream (3' direction) coding sequence. A promoter of a DNA construct, including an oligonucleotide sequence according to the present invention may be linked to a heterologous gene when the presence of the promoter influences transcription from the heterologous gene, including genes for reporter sequences such as luciferase, chloramphenicol acetyl transferase,  $\beta$ -galactosidase and secreted placental alkaline phosphatase.

In a preferred embodiment, the assay is conducted in a whole cell system that has an intracellular receptor which is the target of the screened agent, a promoter or a portion of a promoter with a rel site and a rel-like protein or other transcription protein that binds to the rel site; wherein the intracellular receptor modulates the binding of the rel-like protein or the transcription protein to the rel site. The intracellular receptor, the promoter or a portion of the promoter, or the protein that binds to the rel site may either be endogenous to the cell or transfected into the cell.

In another preferred embodiment, the assay is conducted in an extract of cell having an intracellular receptor, a promoter or a portion of a promoter, with a rel site and a rel-like protein or other protein that binds to the rel site; wherein the intracellular receptor modulates the binding of the rel-like protein or the transcription protein to the rel site.

The binding of the rel-like protein or other transcription protein to the rel site may be measured by techniques known to those skilled in the art, including, but not limited to, mobility shift assay, co-transfection assay, and expression of a reporter gene linked to the promoter.

In a further preferred embodiment, the promoter is activated by an effector, including, but not limited to, tumor necrosis factor, interleukin-1, viruses,

endotoxins, phorbol esters, epidermal growth factor, leukemia inhibitor factor and cAMP agonists.

By "effector" is meant an agent that stimulates the expression of a cytokine to a measurable level. An effector may be endogenously produced in a cell or exogenously added to a cell

In another further preferred embodiment, the claimed assay is conducted in a system including an estrogen receptor, an interleukin 6 promoter or a portion of an IL-6 promoter and NF $\kappa$ B; wherein ER modulates the binding of NF $\kappa$ B or related proteins to the NF $\kappa$ B site on the IL-6 promoter.

The agents discovered by the above assay may either interact directly with an intracellular receptor, or modulate the interaction of a ligand with the intracellular receptor. Thus, in an even further preferred embodiment, a ligand for the intracellular receptor is included in the assay.

While steroids and steroid analogs may exemplify agents identified by the present invention, Applicant is particularly interested in the identification of agents of low molecular weight (less than 10,000 daltons, preferably less than 5,000, and most preferably less than 1,000) which can be readily formulated as useful therapeutic agents.

Such agents can then be screened to ensure that they are specific to tissues with cytokine inflicted pathological conditions with little or no effect on healthy tissues such that the agents can be used in a therapeutic or prophylactic manner. If such agents have some effect on healthy tissues they may still be useful in therapeutic treatment, particularly in those diseases which are life threatening, such as Kaposi's sarcoma or multiple myeloma.

Once isolated, a candidate agent can be put in pharmaceutically acceptable formulations, and used for

specific treatment of diseases and pathological conditions with little or no effect on healthy tissues.

Other features and advantages of the invention will be apparent from the following detailed description of the invention, and from the claims.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows IL-1 and TNF $\alpha$  induced complex formation on the proximal IL-6 promoter.

Figure 2 shows that several distinct NF $\kappa$ B-related complexes induced by IL-1 and TNF $\alpha$  are modulated by estrogen.

Figure 3 shows the effects of estrogen agonist and antagonist, and inhibitors of protein synthesis and protein kinase C on the formation of NF $\kappa$ B-related complexes.

Figure 4 shows the binding characteristics of proteins in NF $\kappa$ B-related complexes with NF $\kappa$ B oligonucleotides.

Figure 5 shows NF $\kappa$ B related proteins in complexes A, B, and C.

#### DETAILED DESCRIPTION OF THE INVENTION

A number of cytokines, including IL-6, IL-8 and IL-11, have related biological effects, i.e., effects on cellular defense in response to infection by stimulating the immune and the acute-phase response and on bone metabolism by increasing bone resorption. Aberrant expression of any of these cytokines results in similar pathological conditions, e.g., all cytokines listed are involved in septic shock. In another example, excessive production IL-8, like IL-6, may be involved in the pathogenesis of several types of inflammatory reactions, particularly neutrophil-dependent tissue damages. These cytokines have similar promoter structures, e.g., their promoters contain binding sites for NF $\kappa$ B or other rel

proteins. It is therefore likely that not only IL-6 but also the other cytokines mentioned above can be targeted by drugs that modulate the binding of NF $\kappa$ B or other rel proteins to their promoter sites through the intracellular receptors.

#### Interleukin 6 and Diseases

Interleukin-6 (IL-6) is a pleiotropic cytokine that is secreted by many different cells, including monocytes, macrophages, certain B-lymphocytes and T-lymphocytes, glial cells, fibroblasts, osteoblasts, and stromal cells (reviewed in references Hirano, T., (1992) "The biology of interleukin-6," Chem. Immunol. 51:153-180.; Kishimoto, T. (1989) "The biology of interleukin-6," Blood 74:1-10.; Kishimoto, T., M. Hibi, M. Murakami, M. Narazaki, M. Saito, and T. Taga (1992) "The molecular biology of interleukin 6 and its receptor," Ciba Found. Symp. 167:5-16; discussion 16-23; and Wolvekamp, M. C., and R. L. Marquet (1990) "Interleukin-6: historical background, genetics and biological significance," Immunol. Lett. 24:1-9). Due to its induction in response to tissue injury, inflammation and infection, IL-6 function is mainly associated with the host's immune and acute phase responses.

IL-6 is an important mediator of intercellular communication not only under pathological conditions but also under normal physiological conditions. It is involved in neural differentiation (Sato, T., S. Nakamura, T. Taga, T. Matsuda, T. Hirano, T. Kishimoto, and Y. Kaziro (1988) "Induction of neuronal differentiation in PC12 cells by B-cell stimulatory factor 2/interleukin 6," Mol. Cell Biol. 8:3546-3549), and proliferation and differentiation during hematopoiesis (Ikebuchi, K., G.G. Wong, S.C. Clark, J.N. Ihle, Y. Hirai, and M. Ogawa (1987) "Interleukin 6 enhancement of interleukin 3-dependent proliferation of

multipotential hemopoietic progenitors," Proc. Natl. Acad. Sci. U.S.A. 84:9035-9039). However, elevated IL-6 expression is usually associated with disease (Yu, X.P., T. Bellido, N. Rice, and S.C. Manolagas (1993)).

5 IL-6 expression is tightly controlled by other factors. Depending on the particular cell type, it can be activated by various stimuli, including tumor necrosis factor (TNF $\alpha$ ) and interleukin-1 (IL-1), viruses, endotoxin (lipopolysaccharides), phorbol  
10 esters, epidermal growth factor (EGF), leukemia inhibitor factor (LIF), and cAMP agonists.

These effectors exhibit their activity through transcriptional effects on the IL-6 promoter as shown by transfection studies (Gruss, H.J., M.A. Brach, and F.  
15 Herrmann (1992) "Involvement of nuclear factor-kappa B in induction of the interleukin-6 gene by leukemia inhibitory factor," Blood 80:2563-2570; Ray, A., S.B. Tatter, L.T. May, and P.B. Sehgal (1988) "Activation of the human "beta 2-interferon/hepatocyte-stimulating  
20 factor/interleukin 6" promoter by cytokines, viruses, and second messenger agonists," Proc. Natl. Acad. Sci. U.S.A. 85:6701-6705). By sequence comparison several potential transcriptional control elements have been identified in the IL-6 promoter, including a cAMP  
25 response element, an AP-1 binding site, and binding elements for the transcription factors NF-IL6 (C/EBPB, LAP, AGP/EBP) and NF $\kappa$ B (Isshiki, H., S. Akira, O. Tanabe, T. Nakajima, T. Shimamoto, T. Hirano, and T. Kishimoto (1990) "Constitutive and interleukin-1 (IL-1)-  
30 inducible factors interact with the IL-1-responsive element in the IL-6 gene," Mol. Cell Biol. 10:2757-2764).

Direct binding of NF-IL6 and NF $\kappa$ B to the IL-6 promoter has been established (Akira, S., H. Isshiki,  
35 T. Sugita, O. Tanabe, S. Kinoshita, Y. Nishio, T. Nakajima, T. Hirano, and T. Kishimoto (1990) "A

nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family," EMBO J. 9:1897-1906; Libermann, T.A. and D. Baltimore, (1990) "Activation of interleukin-6 gene expression through the NF-kappa B transcription factor," Mol. Cell Biol. 10:2327-2334). NF-IL6 belongs to the C/EBP family of leucine zipper proteins. It is induced by IL-1, IL-6 and lipopolysaccharide (LPS), and has been shown to interact with its binding site on the IL-6 promoter and to activate IL-6 expression (Akira, S., H. Isshiki, T. Sugita, O. Tanabe, S. Kinoshita, Y. Nishio, T. Nakajima, T. Hirano, and T. Kishimoto, (1990) "A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family," EMBO J. 9:1897-1906; Chang, C.J., T.T. Chen, H.Y. Lei, D.S. Chen, and S.C. Lee (1990), "Molecular cloning of a transcription factor, AGP/EBP, that belongs to members of the C/EBP family," Mol. Cell Biol. 10:6642-6653; Descombes, P., M. Chojkier, S. Lichtsteiner, E. Falvey, and U. Schibler (1990) "LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein," Genes Dev. 4:1541-1551; Descombes, P., M. Chojkier, S. Lichtsteiner, E. Falvey, and U. Schibler (1990) "LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein," Genes Dev. 4:1541-1551; Poli, V., F.P. Mancini, and R. Cortese (1990) "IL-6DBP, a nuclear protein involved in interleukin-6 signal transduction, defines a new family of leucine zipper proteins related to C/EBP IL-6DBP, a nuclear protein involved in interleukin-6 signal transduction, defines a new family of leucine zipper proteins related to C/EBP," Cell 63:643-653). NFkB is a transcription factor that was originally identified as a heterodimeric complex consisting of a 50 kD protein (p50) and a 65 kD protein (p65) that binds an element in the immunoglobulin kappa light chain enhancer. Both proteins reveal a high

homology to the *Drosophila* morphogen *dorsal* and to the c-rel proto-oncogeny product. The p65 subunit is also functionally related to c-rel (reviewed in references Baeuerle, P. A. (1991) "The inducible transcription activator NF-kappa B: regulation by distinct protein subunits" Biochim. Biophys. Acta 1072:63-80; Blank, V., P. Kourilsky, and A. Israel (1992) "NF-kappa B and related proteins: Rel/dorsal homologues meet ankyrin-like repeats," Trends. Biochem. Sci. 17:135-140; and Liou, H.C. and D. Baltimore (1993) "Regulation of the NF-kappa B/rel transcription factor and I kappa B inhibitor system," Curr. Opin. Cell Biol. 5:477-487). Recently, additional proteins (p49/p52 and relB/p68) have been identified that are functionally related to p50 and p65 (Henkel, T., T. Machleidt, I. Alkalay, M. Kronke, Y. Ben-Neriah, and P.A. Baeuerle (1993) "Rapid proteolysis of I kappa B-alpha is necessary for activation of transcription factor NF-kappa B," Nature 365:182-185; Perkins, N.D., R.M. Schmid, C.S. Duckett, K. Leung, N.R. Rice, and G.J. Nabel (1992) "Distinct combinations of NF-kappa B subunits determine the specificity of transcriptional activation," Proc. Natl. Acad. Sci. U.S.A. 89:1529-1533; Ryseck, R.P., P. Bull, M. Takamiya, V. Bours, U. Siebenlist, P. Dobrzanski, and R. Bravo (1992) "RelB, a new Rel family transcription activator that can interact with p50-NF-kappa B," Mol. Cell Biol. 12:674-684; Ryseck, R.P., P. Bull, M. Takamiya, V. Bours, U. Siebenlist, P. Dobrzanski, and R. Bravo (1992) "RelB, a new Rel family transcription activator that can interact with p50-NF-kappa B," Mol. Cell Biol. 12:674-684; Schmid, R.M., N.D. Perkins, C.S. Duckett, P.C. Andrews, and G.J. Nabel (1991) "Cloning of an NF-kappa B subunit which stimulates HIV transcription in synergy with p65," Nature 352:733-736). NFkB is located in the cytosol complexes with an inhibitory protein of the IkB family. Upon induction, NFkB



dissociates from I $\kappa$ B and translocates into the nucleus where it binds and activates specific promoters (Baeuerle, P.A. and D. Baltimore (1988) "I kappa B: a specific inhibitor of the NF-kappa B transcription factor," Science 242:540-546; Ghosh, S. and D. Baltimore (1990) "Activation in vitro of NF-kappa B by phosphorylation of its inhibitor I kappa B," Nature 344:678-682). Binding of NF $\kappa$ B-like factors to the consensus site of the IL-6 promoter is induced by IL-1, TNF $\alpha$ , LIF, LPS and phorbol esters, varying with the particular cell type (Gruss, H.J., M.A. Brach, and F. Herrmann (1992) "Involvement of nuclear factor-kappa B in induction of the interleukin-6 gene by leukemia inhibitory factor," Blood 80:2563-2570; Libermann, T.A. and D. Baltimore (1990) "Activation of interleukin-6 gene expression through the NF-kappa B transcription factor," Mol. Cell Biol. 10:2327-2334; Shimizu, H., K. Mitomo, T. Watanabe, S. Okamoto, and K. Yamamoto (1990) "Involvement of a NF-kappa B-like transcription factor in the activation of the interleukin-6 gene by inflammatory lymphokines," Mol. Cell Biol. 10:561-568; Zhang, Y.H., J.X. Lin, and J. Vilcek (1990) "Interleukin-6 induction by tumor necrosis factor and interleukin-1 in human fibroblasts involves activation of a nuclear factor binding to a kappa B-like sequence," Mol. Cell Biol. 10:3818-3823).

Unregulated expression of IL-6 is linked to a number of diseases (Bauer, J. and F. Herrmann (1991) "Interleukin-6 in clinical medicine," Ann. Hematol. 62:203-210; Hirano, T. (1992) "Interleukin-6 and its relation to inflammation and disease," Clin. Immunol. Immunopathol. 62:S60-S65) such as postmenopausal osteoporosis after loss of ovarian function (Roodman, G.D. (1992) "Interleukin-6: an osteotropic factor?" J. Bone Miner. Res. 7:475-478). Ex vivo cultures of bone marrow from ovariectomized mice show an increase of

osteoclastogenesis compared with cultures from sham-operated animals. This increase in osteoclast development can be prevented by injection of an anti-IL-6 antibody or by administration of estrogen (Jilka, R.L., G. Hangoc, G. Girasole, G. Passeri, D.C. Williams, J.S. Abrams, B. Boyce, H. Broxmeyer, and S.C. Manolagas (1992) "Increased osteoclast development after estrogen loss: mediation by interleukin-6," Science 257:88-91). In mice that carry a null mutation for IL-6, ovariectomy does not affect bone volume or osteoclast number as seen with normal mice (Balena, R., F. Costantini, M. Yamamoto, A. Markatos, R. Cortese, G.A. Rodan, and V. Poli (1993) "Mice with IL-6 gene knock-out do not lose cancellous bone after ovariectomy," J. Bone Miner. Res. 8:S130 [Abstract]).

#### Regulation of Interleukin 6 by Estrogen

Estrogen has been found to inhibit IL-6 expression in bone-derived stromal cell lines and osteoblastic cells from rats and mice as well as in nontransformed human bone cells (Girasole, G., R.L. Jilka, G. Passeri, S. Boswell, G. Boder, D.C. Williams, and S.C. Manolagas (1992) "17 beta-estradiol inhibits interleukin-6 production by bone marrow-derived stromal cells and osteoblasts in vitro: a potential mechanism for the antiosteoporotic effect of estrogens," J. Clin. Invest. 89:883-891). This effect of estrogen on IL-6 expression is not restricted to bone tissue but has also been shown for uterine cells (Jacobs, A.L., P.B. Sehgal, J. Julian, and D.D. Carson (1992) "Secretion and hormonal regulation of interleukin-6 production by mouse uterine stromal and polarized epithelial cells cultured in vitro," Endocrinology 131:1037-1046; Tabibzadeh, S.S., U. Santhanam, P.B. Sehgal, and L.T. May (1989) "Cytokine-induced production of IFN-beta 2/IL-6 by freshly explanted human endometrial stromal cells. Modulation by estradiol-17 beta," J. Immunol. 142:3134-

3139). There are only a few other genes known to be negatively regulated by estrogen agonists (Adler, S., M.L. Waterman, X. He, and M.G. Rosenfeld (1988) "Steroid receptor-mediated inhibition of rat prolactin gene expression does not require the receptor DNA-binding domain," Cell 52:685-695; Ree, A.H., B.F. Landmark, W. Eskild, F.O. Levy, H. Lahooti, T. Jahnsen, A. Aakvaag, and V. Hansson (1989) "Autologous down-regulation of messenger ribonucleic acid and protein levels for estrogen receptors in MCF-7 cells: an inverse correlation to progesterone receptor levels," Endocrinology 124:2577-2583).

To investigate the mechanism of the estrogen effect, Applicant performed a series of DNA-binding experiments using the human IL-6 promoter. Co-transfection studies showed that the proximal 225 bps of the IL-6 promoter mediate both the induction of the reporter gene by IL-1 and TNF $\alpha$  as well as the repression by estradiol. The repression by estradiol also required the expression of the estrogen receptor (ER).

Using gel retardation assays, no specific binding of the ER to the proximal 225 bp could be detected. However, nuclear extracts from +/-LDA11 bone marrow stromal cells that revealed IL-6 regulation by IL-1, TNF $\alpha$ , and estradiol showed an induced complex with a -225 to -52 promoter fragment when the cell were treated with IL-1 and TNF $\alpha$ . Induction of the complex was fast (10 minutes) but transient. Pretreatment of the cells with estradiol increased the intensity as well as the mobility of the complex.

To identify the proteins involved in the formation of the complex, antibody supershift experiments were carried out using antibodies against factors with potential binding sites in this promoter fragment including c-jun, NF-IL6, c-rel, and NF $\kappa$ B p50 and p65

proteins. Only anti-p50 and anti-p65 had an effect and abolished the formation of the induced complex.

An oligonucleotide covering the potential NF $\kappa$ B site of the IL-6 promoter competed against the induced binding to this fragment, while an oligonucleotide covering the NF-IL6 site was ineffective. When the NF $\kappa$ B oligonucleotide was used as probe, three IL-1/TNF $\alpha$ -induced complexes were observed.

Pretreatment with estradiol decreased the intensity of the slowest complex and strongly increased the intensity of the fastest migrating complex. The three bands were differentially supershifted (i.e., further decrease in the mobility of the complex due to binding of the antibody) by anti-p50 and anti-p65 antibodies, while none of several other antibodies tested, including anti-ER antibody, had any effect. Methylation interference assays showed identical DNA contact sites for all three complexes.

Ray, et al., J. Biol. Chem., 269(17):12940-946 (1994), not admitted to be prior art, describe that activation of the IL-6 promoter, elicited by a combination of NF-IL6 and the p65 subunit of NF $\kappa$ B, can be inhibited by the wt ER but not by an ER containing a mutation in its DNA binding domain. Furthermore, the repression of the IL-6 promoter by a combination ER and 17 $\beta$ -estradiol did not appear to be mediated via high affinity binding of the receptor to the promoter.

These data suggest that negative regulation by estrogen is mediated through the IL-6 promoter and is estrogen receptor dependent. Inhibition of IL-6 expression by estrogen is mediated through control of the transcriptional activity of NF $\kappa$ B or closely related proteins on the IL-6 promoter.

Mukaida, et al., J. Biol. Chem., 269(18):13289-295 (1994), not admitted to be prior art, describe that a glucocorticoid, dexamethasone, inhibited IL-8 production

at the transcriptional level. Mutation of either the AP-1 or NF-IL6 binding site on the IL-8 promoter did not abolish IL-8 gene repression by dexamethasone, suggesting that these sites were not targets for dexamethasone. Yet dexamethasone diminished the IL-1 induced formation of NF $\kappa$ B complexes.

The invention will now be described in greater detail by reference to the following non-limiting examples regarding the regulation of interleukin 6 transcription by estrogen receptor.

#### Examples

A candidate agent will be screened by either A) direct evaluation of protein binding to rel-sites, or B) indirect evaluation of binding to rel-sites.

##### A) Direct evaluation of protein binding to rel-sites

Cells selected for expression of the necessary components will be treated with the agent or vehicle control and an inducer (e.g., phorbol ester, cytokines, lipopolysaccharides). Cellular extracts prepared from those cells (e.g., whole cell, cytosolic, or nuclear extracts) will be analyzed for their DNA-binding using cytokine promoter fragments or various rel-sites as probes. Binding will be analyzed qualitatively (i.e., comparing pattern) and quantitatively comparing extracts from cells treated with vehicle or the agent.

##### B) Indirect evaluation of binding to rel-sites by

###### 1) Measuring endogenous cytokine expression.

Cells selected for expression of the necessary components and their production of cytokine will be treated with the agent or vehicle control and an inducer (phorbol ester, cytokines, lipopolysaccharides). Activity of the agent will be quantitatively assessed by

measuring of cytokine using standard assays known to those skilled in the art.

2) Measuring the expression of a reporter introduced into the cell.

5 By means of transfection a reporter construct will be introduced into the cells that expresses an easily measurable protein under the control of a cytokine promoter or fragments thereof or isolated rel-sites. The other necessary components are either expressed  
10 endogenously by the cells or provided by cotransfection of expression vectors for the particular component. Cells will be treated with the agent or vehicle control and an effector (phorbol ester, cytokines, lipopolysaccharides). The activity of the agent will be  
15 analyzed quantitatively by measuring the expression of the reporter protein.

Agents will also be tested for their binding to IRs by traditional binding assays as well as for their activity to effect the classical mechanism of gene  
20 regulation by IRs. An agent that binds to IRs and regulates binding of rel proteins to cytokine promoters but does not activate the classical mechanism of IR action is a potential drug candidate for the specific treatment of diseases associated with aberrant  
25 expression of cytokines.

Experimental procedures employed in the examples described herein are set forth below:

Transient transfections and mammalian expression constructs

30 Construction of the pERE-tk-Luc reporter plasmid and the vector expressing ER<sub>gly</sub> (pRShER) has been described (Tzukerman, M., A. Esty, D. Santiso-Mere, P. Danielian, M.G. Parker, R.B. Stein, J.W. Pike, and D.P. McDonnell  
35 (1994) "Human estrogen receptor transactivational capacity is determined by both cellular and promoter

context and mediated by two functionally distinct intramolecular regions," Mol. Endocrinol. 8:21-30, incorporated by reference herein). The pIL6[-225]Luc reporter construct was derived from the parental pIL6[-1200]Luc by excision of a NheI-BamHI fragment and religation of the vector fragment after blunt ending with Klenow DNA-polymerase. The parental pIL6[-1200]Luc was constructed by cloning the 1.2 kb IL-6 promoter insert excised with BamHI and KpnI from pCAT-M54-IL6(-) into the corresponding sites of the luciferase vector Lucpl.

C3H10T1/2 cells were seeded in phenol-red-free DMEM supplemented with 10% FBS at 80,000 cells per well (12-well plates). The cells were transfected by calcium phosphate precipitation (Peterson, J.L. and O.W. McBride (1980) "Cotransfer of linked eukaryotic genes and efficient transfer of hypoxanthine phosphoribosyltransferase by DNA-mediated gene transfer," Proc. Natl. Acad. Sci. U.S.A. 77:1583-1587) with 0.5 mg pIL6[-225]Luc alone or together with 0.05 mg pRShER or 0.1 mg HE0 using pGEM as carrier to adjust to 2 mg total DNA in the transfection mix. After 4 h at 37 °C the cells were treated with 7% DMSO for 30 min followed by a medium change and addition of hormones. The following day the cells were induced with TNF $\alpha$  and IL-1b (1 nM each) for 24 h. After a brief wash with PBS the cells were lysed in 200  $\mu$ l lysis buffer (25 mM Tris [pH 7.8], 2 mM DTT, 2 mM 1,2-diaminocyclohexane-N,N,N',N'-tetraacetic acid, 10% glycerol, 1% Triton X-100). To 20  $\mu$ l of each extract 100  $\mu$ l of reagent (20 mM Tricine [pH 7.8], 1.07 mM (MgCO<sub>3</sub>)<sub>4</sub>Mg(OH)<sub>2</sub>, 2.67 mM MgSO<sub>4</sub>, 0.1 mM EDTA, 33.3 mM DTT, 279 mM coenzyme A, 470 mM luciferin, 530 mM ATP) was added and luciferase activity was measured immediately with a Dynatech luminometer in cycle mode.

Antibodies, IL-6 ELISA, and ER assay

Peptides used to raise the following antibodies in rabbits correspond to amino acid residues 91-105 of murine c-jun (Ryder and Nathans (1988) "Induction of protooncogene c-jun by serum growth factors," Proc. Natl. Acad. Sci. USA 85:8464-8467), 278-296 of murine NF-IL6 (Chang, C.J., T.T. Chen, H.Y. Lei, D.S. Chen, and S.C. Lee (1990) "Molecular cloning of a transcription factor, AGP/EBP, that belongs to members of the C/EBP family," Mol. Cell Biol. 10:6642-6653), 152-176 of murine c-rel (Bull, P., K.L. Morley, M.F. Hoekstra, T. Hunter, and I.M. Verma (1990) "The mouse c-rel protein has an N-terminal regulatory domain and a C-terminal transcriptional transactivation domain," Mol. Cell Biol. 10:5473-5485; Inoue, J., L.D. Kerr, L.J. Ransone, E. Bengal, T. Hunter, and I.M. Verma (1991) "c-rel activates but v-rel suppresses transcription from kappa B sites," Proc. Natl. Acad. Sci. U.S.A. 88:3715-3719), 347-361 of murine p50 (Ghosh, S., A.M. Gifford, L.R. Riviere, P. Tempst, G.P. Nolan, and D. Baltimore (1990) "Cloning of the p50 DNA binding subunit of NF-kappa B: homology to rel and dorsal," Cell 62:1019-1029), and 3-19 of human p65 (88% homology with murine p65) (Nolan, G.P., S. Ghosh, H.C. Liou, P. Tempst, and D. Baltimore (1991) "DNA binding and I kappa B inhibition of the cloned p65 subunit of NF-kappa B, a rel-related polypeptide," Cell 64:961-969; Ruben, S.M., P.J. Dillon, R. Schreck, T. Henkel, C.H. Chen, M. Maher, P.A. Baeuerle, and C.A. Rosen (1991) "Isolation of a rel-related human cDNA that potentially encodes the 65-kD subunit of NF-kappa B [letter]," Science 254:11). All the references mentioned above are incorporated by reference herein. All antibodies listed above were obtained affinity purified at a concentration of 1 mg/ml from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA). Anti-TBP was a protein-A purified serum preparation from



a rabbit immunized with the full length human recombinant protein and reported to react with TBP from mouse, rat, and human origin (Santa Cruz Biotechnology, Inc.). Anti-ER is a mouse monoclonal antibody (IgG2a) raised against a peptide corresponding to amino acid residues 8-22 of the murine ER. IL-6 concentration in tissue culture supernatants was determined by use of an IL-6 ELISA kit (Endogen, Inc., Boston, MA) using murine IL-6 as standard.

ER in +/-LDA11 cells was measured in whole cell extracts. After washing and counting, cells were homogenized in buffer containing 50 mM Tris [pH 7.5], 30% glycerol, 500 mM KCl, 1 mM EDTA, 1 mM PMSF, and 5 mM DTT. After 30 min on ice the homogenate was centrifuged (100,000 g, 4 °C, 1 h). The supernatant was taken as whole cell extract, adjusted to 0.5% CHAPS, and incubated with 5 nM [<sup>3</sup>H]estradiol in the absence or presence of a 200-fold excess of DES overnight at 4 °C. After incubation with anti-ER antibody, the complexes formed were precipitated with protein-A sepharose (Pharmacia), washed three times with 10 mM Tris [pH 7.5]/0.5% CHAPS, and measured by liquid scintillation counting.

#### Electrophoretic mobility shift assay (EMSA) and methylation interference assay

DNA binding studies were carried out with nuclear extracts from +/-LDA11 cells, extracts from yeast expressing recombinant human ER<sub>gly</sub>, and purified p50 and p49 proteins. +/-LDA11 cells were maintained under conditions as described (Girasole, G., R.L. Jilka, G. Passeri, S. Boswell, G. Boder, D.C. Williams, and S.C. Manolagas (1992) "17 beta-estradiol inhibits interleukin-6 production by bone marrow-derived stromal cells and osteoblasts in vitro: a potential mechanism for the antiosteoporotic effect of estrogens,"

J. Clin. Invest. 89:883-891). To prepare nuclear extracts the cells were seeded in phenol-red-free McCoy's medium supplemented with 10% FBS and pretreated with hormone for 24 h if not indicated otherwise. After  
5 adjusting the medium to 2% FBS, the cells were induced with TNF $\alpha$  and IL-1b (1 nM each) for varying periods. In cases where cycloheximide (10 mg/ml) or the kinase inhibitor H7 (50 mM) were included, those compounds were added 5 min before induction. Incubation was stopped by  
10 two washes with ice cold PBS and cells were lysed in situ in cold buffer A (10 mM HEPES [pH 7.9], 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 0.5 mM DTT, 0.2% Nonidet P-40). Lysates were transferred into microfuge tubes, nuclei pelleted (8000 rpm, 1 min) and resuspended in buffer C  
15 (20 mM HEPES [pH 7.9], 1.5 mM MgCl<sub>2</sub>, 420 mM NaCl, 25% glycerol, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF). After 40 min rocking at 4 °C, samples were centrifuged (15,000 rpm, 10 min) and supernatants taken as nuclear extracts. Bradford protein assays (Bradford, M.M. (1976) "A rapid  
20 and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding," Anal. Biochem. 72:248-254) showed only minimal variations in protein concentrations which did not correlate with hormone or cytokine treatment. Extracts of yeast recombinantly expressing ER<sub>gly</sub> were  
25 prepared from the BJ2168 strain transformed with YE<sub>p</sub>E10 as described (Tzukerman, M., A. Esty, D. Santiso-Mere, P. Danielian, M.G. Parker, R.B. Stein, J.W. Pike, and D.P. McDonnell (1994) "Human estrogen receptor  
30 transactivational capacity is determined by both cellular and promoter context and mediated by two functionally distinct intramolecular regions," Mol. Endocrinol. 8:21-30). Purified, *Escherichia coli* expressed human p50 and p49 proteins were purchased from  
35 Promega (Madison, WI).

For EMSA, 2ml of the extracts were preincubated with 2 mg poly[dI-dC] in binding buffer adjusted to 20 mM HEPES [pH 7.9], 40 mM NaCl, 20 mM KCl, 2.5 mM MgCl<sub>2</sub>, 10% glycerol, 0.1 mg/ml BSA, and 1 mM DTT. When the -225 to -52 IL-6 promoter fragment was used as probe 0.5 mg of Bluescript plasmid (Stratagene, La Jolla, CA) was also included. After 20 min on ice, the probe was added and the incubation continued for 20 min at room temperature. When antibodies were included, 1 mg was added 20 min after the probe and the incubation continued for 40 min at 4 °C. The complexes formed were analyzed on non-denaturing polyacrylamide gels (4% acrylamide/0.05% BIS; 2x200 mm) at 4 °C and 15 V/cm in 0.25xTBE. Probes were either double stranded oligonucleotides corresponding to the regions -82 to -47 (ATCAATGTGGGATTTTCCCATGAGTCTCAATATTA) and -172 to -131 (CTAAAGGACGTCACATTGCACAATCTTAATAAGGTTTCCAAT) of the human IL-6 promoter and to the ERE of the vitellogenin promoter (Tzukerman, M., X.K. Zhang, T. Hermann, K.N. Wills, G. Graupner, and M. Pfahl (1990) "The human estrogen receptor has transcriptional activator and repressor functions in the absence of ligand," New Biol. 2:613-620) or the -225 to -52 NheI-SspI IL-6 promoter fragment. All probes were either labeled with [ $\gamma^{32}$ P]ATP using T4-polynucleotide kinase or with [ $\alpha^{32}$ P]dATP using Klenow polymerase and subsequently purified by polyacrylamide gel electrophoresis.

For methylation interference assays, the -82 to -47 probe labeled with [ $\gamma^{32}$ P]ATP either on the upper or the lower strand was subjected to limited DMS-methylation (Maxam, A.M. and W. Gilbert (1980) "Sequencing end-labeled DNA with base-specific chemical cleavages," Methods Enzymol. 65:499-560). EMSA was performed as described above, scaled up 10-fold. Gels were blotted onto NA45 anion exchange membranes (Schleicher & Schuell) in 0.5xTBE for 30 min at 30 V (Singh, H., J.H.

LeBowitz, A.S. Baldwin, Jr., and P.A. Sharp (1988).  
"Molecular cloning of an enhancer binding protein:  
isolation by screening of an expression library with a  
recognition site DNA," Cell 52:415-423). After  
5 autoradiography, the DNA corresponding to the various  
complexes and the unretarded probe was eluted (10 min at  
65 °C in 20 mM Tris [pH 8.0], 1 M NaCl, 0.1 mM EDTA) and  
purified by phenol/chloroform extraction and ethanol  
precipitation. After strand cleavage in 1 M piperidine  
10 (30 min at 90 °C) the fragments were resolved on  
denaturing polyacrylamide gels (12% acrylamide/0.6%  
BIS).

Example 1. Screening for ER mediated inhibition of IL-6  
15 promoter activity

It has been shown that IL-6 repression is regulated  
by estradiol at the mRNA level (Girasole, G., R.L.  
Jilka, G. Passeri, S. Boswell, G. Boder, D.C. Williams,  
and S.C. Manolagas (1992) "17 beta-estradiol inhibits  
20 interleukin-6 production by bone marrow-derived stromal  
cells and osteoblasts in vitro: a potential mechanism  
for the antiosteoporotic effect of estrogens," J. Clin.  
Invest. 89:883-891; Jacobs, A.L., P.B. Sehgal, J.  
Julian, and D.D. Carson (1992) "Secretion and hormonal  
25 regulation of interleukin-6 production by mouse uterine  
stromal and polarized epithelial cells cultured in  
vitro," Endocrinology 131:1037-1046). To determine if  
estrogen or a candidate agent acts directly on IL-6  
transcription, we transfected a reporter construct,  
30 expressing the firefly luciferase under the control of  
the human IL-6 promoter region from -225 to +14, into  
the murine fibroblast cell line C3H10T1/2. These cells  
can be considered as pre-osteoblasts since they  
differentiate into osteogenic cells in response to bone  
35 morphogenic protein-2 (Katagiri, T., A. Yamaguchi, T.  
Ikeda, S. Yoshiki, J.M. Wozney, V. Rosen, E.A. Wang, H.

Tanaka, S. Omura, and T. Suda (1990) "The non-osteogenic mouse pluripotent cell line, C3H10T1/2, is induced to differentiate into osteoblastic cells by recombinant human bone morphogenetic protein-2," Biochem. Biophys. Res. Commun. 172:295-299).

Therefore, C3H10T1/2 cells were transfected with a luciferase expression vector under the control of the proximal human IL-6 promoter (pIL6[-225]Luc) alone or together with the expression vector for the wild-type human ER<sub>gly</sub> (pRShER). After pretreatment with varying concentrations of estradiol for 24 hours, the cultures were induced with 1 nM each of TNF $\alpha$  and IL-1 or left uninduced and 24 h later cells were harvested and extracts analyzed for luciferase activity.

Treatment of transfected cells with IL-1 and TNF $\alpha$  induced a 5-fold increase in luciferase activity over basal levels. Without cotransfection of a plasmid expressing the estrogen receptor, treatment with estradiol had no effect. However, with the expression of estrogen receptor by cotransfection, treatment with estradiol resulted in a strong, dose-dependent repression of luciferase activity.

Repression was observed with the wild-type human ER (ER<sub>gly</sub>) as well as with an ER variant containing a glycine to valine point mutation in the hormone binding domain (ER<sub>val</sub>) (Tora, L., A. Mullick, D. Metzger, M. Ponglikitmongkol, I. Park, and P. Chambon (1989) "The cloned human oestrogen receptor contains a mutation which alters its hormone binding properties," EMBO J. 8:1981-1986). While ER<sub>val</sub> required a higher estradiol concentration, it exhibited a stronger repression. This is consistent with the finding that in induction experiments ER<sub>gly</sub> responds at lower hormone concentrations but has considerable basal activity (Tzukerman, M., X.K. Zhang, T. Hermann, K.N. Wills, G. Graupner, and M. Pfahl (1990) "The human estrogen receptor has transcriptional

activator and repressor functions in the absence of ligand," New Biol. 2:613-620).

The dependence of the estrogen effect on cotransfected ER suggested that C3H10T1/2 cells do not express functional endogenous ER. This was confirmed by transfecting the cell with a luciferase reporter under the control of the vitellogenin estrogen response element (ERE) (Klein-Hitpass, L., M. Schorpp, U. Wagner, and G.U. Ryffel (1986) "An estrogen-responsive element derived from the 5' flanking region of the Xenopus vitellogenin A2 gene functions in transfected human cells," Cell 46:1053-1061).

Therefore, C3H10T1/2 cells were transfected with a luciferase expression vector under the control of the minimal thymidine kinase promoter and the vitellogenin estrogen response element (pERE-tk-Luc) alone or together with pRShER. 24 h after treatment with 10 nM estradiol or vehicle cells were harvested and extracts analyzed for luciferase activity. Induction of luciferase activity by estradiol was only observed in the presence of cotransfected ER.

In addition, C3H10T1/2 cells were incubated with or without 10 nM estradiol. After 24 h the cultures were induced with TNF $\alpha$  and IL-1 (1 nM each) or left uninduced for additional 24 h. IL-6 in the supernatants was assayed by an ELISA specific for murine IL-6. C3H10T1/2 cells responded to IL-1 and TNF $\alpha$  treatment with strongly increased production of endogenous IL-6, but unlike other osteogenic or stromal cells containing endogenous ER (Girasole, G., R.L. Jilka, G. Passeri, S. Boswell, G. Boder, D.C. Williams, and S.C. Manolagas (1992) "17 beta-estradiol inhibits interleukin-6 production by bone marrow-derived stromal cells and osteoblasts in vitro: a potential mechanism for the antiosteoporotic effect of estrogens," J. Clin. Invest. 89:883-891), IL-6 levels were not decreased by estradiol. These data suggest

that the inhibition of IL-6 expression is at the transcriptional level and mediated through the ER.

By cotransfection studies using the preosteoblastic cell line C3H10T1/2, we showed that IL-1/TNF $\alpha$ -induced

5 activation of the proximal IL-6 promoter region could be inhibited by estrogen. This inhibition was estrogen receptor dependent and was observed with both the wild-type human ER (ER<sub>gly</sub>) and the ER<sub>val</sub> variant. Similar results have been obtained by others in both HeLa cells  
10 cotransfected with ER<sub>val</sub>, and in MBA13 cells, a preosteoblastic cell line expressing endogenous ER. Together with the described effects of estrogen on IL-6 mRNA (Girasole, G., R.L. Jilka, G. Passeri, S. Boswell, G. Boder, D.C. Williams, and S.C. Manolagas (1992) "17  
15 beta-estradiol inhibits interleukin-6 production by bone marrow-derived stromal cells and osteoblasts in vitro: a potential mechanism for the antiosteoporotic effect of estrogens," J. Clin. Invest. 89:883-891; Jacobs, A.L., P.B. Sehgal, J. Julian, and D.D. Carson (1992)  
20 "Secretion and hormonal regulation of interleukin-6 production by mouse uterine stromal and polarized epithelial cells cultured in vitro," Endocrinology 131:1037-1046), these results suggest a transcriptional mechanism of estrogen-induced inhibition.

25 A candidate agent can be screened using the above assay, replacing estradiol with said agent.

Example 2. Screening agents that modulates binding of NF $\kappa$ B related proteins to the proximal IL-6 promoter

30 A cell line that expresses ER (+/+LDA11)

An exemplary assay system is a cell line that expressed all the necessary components endogenously, including the ER. The bone marrow derived murine  
35 stromal cell line +/-LDA11 has been shown to respond to IL-1 and TNF $\alpha$  treatment with strongly increased

secretion of IL-6. Treatment with estradiol inhibits this induction of IL-6 as shown for the protein and its mRNA (Girasole, G., R.L. Jilka, G. Passeri, S. Boswell, G. Boder, D.C. Williams, and S.C. Manolagas (1992) "17 beta-estradiol inhibits interleukin-6 production by bone marrow-derived stromal cells and osteoblasts in vitro: a potential mechanism for the antiosteoporotic effect of estrogens," J. Clin. Invest. 89:883-891).

To verify that ER is actually present in +/+LDA11 cells, hormone binding studies were carried out. Initial experiments showed a low number of specific estradiol binding sites in high salt extracts from these cells. Using a monoclonal antibody directed against the amino terminus of the ER, specifically bound [<sup>3</sup>H]estradiol was immunoprecipitated confirming that the binding sites represented ER. From those studies we calculated that +/+LDA11 cells contain approximately 1000 ER molecules per cell.

However, when using electrophoretic mobility shift assays (EMSA) in combination with the vitellogenin ERE as a probe, ER-specific DNA binding activity could not be detected in nuclear extracts from +/+LDA11 treated with estradiol and/or IL-1 and TNF $\alpha$ . Nuclear extracts of +/+LDA11 cells pretreated with estradiol (10 nM) and TNF $\alpha$  and IL-1 (1nM each for 40 min) as indicated or yeast extract containing recombinantly expressed human wild-type ER<sub>gly</sub> were incubated with the vitellogenin ERE as probe in the absence or presence of anti-ER antibody. Complexes formed were analyzed by EMSA.

The complexes detected are unrelated to the ER since they were not significantly affected by anti-ER antibody. Controls using ER containing extracts obtained from a yeast expression system gave rise to two slowly migrating complexes that were specifically shifted with the anti-ER antibody. These data suggest



that ER is present in +/+LDA11 cells but at concentrations too low to be detected by EMSA.

DNA-binding activity of nuclear extracts to the IL-6 promoter

To study the molecular mechanism of IL-6 induction, and its repression by estrogen, nuclear extracts from +/+LDA11 cells were analyzed for DNA-binding activity to the IL-6 promoter region that mediated cytokine induction and estrogen suppression in the cotransfection experiments. Since this DNA fragment showed a high background binding with nuclear extracts the most proximal region containing the TATA box was removed leaving a fragment from -225 to -52 upstream of the transcriptional start site. This region of the promoter contains consensus binding sites for several transcription factors including a core sequence of the cAMP response element (CRE), a binding site for the leucine zipper protein NF-IL6, and a NF $\kappa$ B site (Isshiki, H., S. Akira, O. Tanabe, T. Nakajima, T. Shimamoto, T. Hirano, and T. Kishimoto (1990) "Constitutive and interleukin-1 (IL-1)-inducible factors interact with the IL-1-responsive element in the IL-6 gene," Mol. Cell Biol. 10:2757-2764).

Binding of NF-IL6 and NF $\kappa$ B-like proteins to these sites has been demonstrated (Akira, S., H. Isshiki, T. Sugita, O. Tanabe, S. Kinoshita, Y. Nishio, T. Nakajima, T. Hirano, and T. Kishimoto (1990) "A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family," EMBO J. 9:1897-1906; Libermann, T.A. and D. Baltimore (1990) "Activation of interleukin-6 gene expression through the NF-kappa B transcription factor," Mol. Cell Biol. 10:2327-2334). This fragment was incubated with nuclear extracts from +/+LDA11 cells that had been treated with IL-1 and TNF $\alpha$  for various times. +/+LDA11 cells were pretreated with 10 nM estradiol as indicated.

After 24 h the cells were induced with TNF $\alpha$  and IL-1 (1 nM each) for various periods of time. Induction was stopped by cell lysis and nuclear extracts were analyzed by EMSA using the -225 to -52 IL-6 promoter fragment as probe (Figure 1a).

Complexes formed were analyzed by EMSA. After treatment with the cytokines an inducible complex was observed. The intensity of the complex was maximal already after 10 min treatment with IL-1 and TNF $\alpha$  and decreased gradually over time. After 2 hours of induction the intensity of the complex was significantly reduced.

Pretreatment of the cells with estradiol had no effect on the binding capacity of extracts from uninduced cells. However, estradiol pretreatment resulted in a marked increase of the induced complex with induction intervals from 10 min to 40 min but only a slight effect on the complex after 2 h of induction. In addition to the increased intensity, pretreatment with estradiol also caused a qualitative change, increasing the mobility of the complex.

#### Detecting the composition of the DNA-binding complex

To investigate the nature of the complex and the proteins potentially involved, we incubated the binding reactions with antibodies directed against several potential binding factors. Nuclear extract from +/-LDA11 cells treated with estradiol (10 nM) and TNF $\alpha$  and IL-1 (1nM each for 10 min) as indicated were incubated with the -225 to -52 probe in the absence or presence of various antibodies. Complexes formed were analyzed by EMSA.

Fig. 1b shows that none of the antibodies tested affected DNA binding of extracts from uninduced cells. Neither anti-c-jun, nor anti-c-rel, nor anti-NF-IL6

antibodies had any effect on the cytokine induced complexes.

However, anti-p50 and anti-p65, antibodies directed against the two proteins in the NF $\kappa$ B complex, abolished formation of the complex (lanes 10-13). This was observed with extracts derived from cells treated or untreated with estradiol, over the whole period of cytokine induction (Fig. 1b depicts the results at 10 minutes after the induction). With longer exposures, a very weak complex of low mobility was detected, probably resulting from a supershift of the induced complex by anti-p50 and anti-p65.

Although ER binding activity to the vitellogenin ERE was not detectable in +/+LDA11 extracts, we tested whether the ER was involved in complex formation on the IL-6 promoter fragment. Nuclear extracts of +/+LDA11 cells treated with estradiol (10 nM) and TNF $\alpha$  and IL-1 (1nM each for 40 min) as indicated or yeast extract containing recombinantly expressed human wild-type ER were incubated with the -225 to -52 probe in the absence or presence of anti-ER antibody. Complexes formed were analyzed by EMSA.

Fig. 1c shows that independent of cytokine induction or estradiol treatment, addition of the anti-ER antibody did not significantly affect any of the complexes, induced or constitutive. In addition, when yeast extracts containing high concentrations of recombinant ER were incubated with the IL-6 promoter fragment no specific binding of the ER was detected (lanes 7 and 8). The weak bands observed are unrelated to the ER, since they were not affected by the anti-ER antibody.

The results from the antibody gel shift experiments were further supported by oligonucleotide competition studies. Nuclear extract from +/+LDA11 cells treated with estradiol (10 nM) and TNF $\alpha$  and IL-1 (1nM each for 10 min) as indicated were incubated with the -225 to -52

probe in the absence or presence of a 400-fold molar excess of oligonucleotides corresponding to the regions of -82 to -47 and -172 to -131 of the human IL-6 promoter. Complexes formed were analyzed by EMSA.

5       The arrows in Fig. 1d indicate the complexes formed upon induction with TNF $\alpha$  and IL-1. Inclusion of an oligonucleotide covering the NF-IL6 site, the CRE, and an adjacent CCAAT-box of the IL-6 promoter (-172 to -131) in the binding reaction in 400-fold excess over the  
10       labeled -225 to -52 fragment did not affect any of the complexes, constitutive or cytokine-induced (lanes 4 and 7). However, an oligonucleotide covering the putative NF $\kappa$ B site and adjacent sequences (-82 to -47) specifically abolished the formation of the cytokine  
15       induced complexes (lanes 3 and 6).

20       The antibody experiments and the oligonucleotide competition studies suggested that IL-1 and TNF $\alpha$  specifically activated NF $\kappa$ B or related proteins. No binding of c-jun (AP-1), NF-IL6, c-rel, or ER was detected.

25       The lack of NF-IL6 binding is surprising, since induction and binding of this transcription factor in response to IL-1 has been reported for other cells (Akira, S., H. Isshiki, T. Sugita, O. Tanabe, S. Kinoshita, Y. Nishio, T. Nakajima, T. Hirano, and T. Kishimoto (1990) "A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family," EMBO J. 9:1897-1906; Inoue, J., L.D. Kerr, L.J. Ransone, E. Bengal, T. Hunter, and I.M. Verma (1991) "c-rel activates but v-rel suppresses transcription from kappa B sites," Proc. Natl. Acad. Sci. USA 88:3715-3719). Our DNA binding  
30       experiments show that in the bone marrow derived +/+LDA11 cell IL-1 and TNF $\alpha$  induce the binding of NF $\kappa$ B or closely related proteins to the IL-6 promoter.  
35       Similar results have been obtained in different cell types (H. Shimizu, K. Mitomo, T. Watanabe, S. Okamoto,

and K. Yamamoto (1990) "Involvement of a NF-kappa B-like transcription factor in the activation of the interleukin-6 gene by inflammatory lymphokines," Mol. Cell Biol. 10:561-568; Zhang, Y.H., J.X. Lin, and J. Vilcek (1990) "Interleukin-6 induction by tumor necrosis factor and interleukin-1 in human fibroblasts involves activation of a nuclear factor binding to a kappa B-like sequence," Mol. Cell Biol. 10:3818-3823). Neither induced nor uninduced binding of several other factors with potential binding sites in the proximal IL-6 promoter fragment could be detected, including AP-1. This transcription factor is one of the paradigms for direct inhibition by intracellular receptor including GR, RAR and TR. The mechanism of AP-1 inhibition can involve protein-protein interaction and/or competition for DNA binding depending on the particular gene (Diamond, M.I., J.N. Miner, S.K. Yoshinaga, and K.R. Yamamoto (1990) "Transcription factor interactions: selectors of positive or negative regulation from a single DNA element," Science 249:1266-1272; Schüle, R., K. Umesono, D.J. Mangelsdorf, J. Bolado, J.W. Pike, and R.M. Evans (1990) "Jun-Fos and receptors for vitamins A and D recognize a common response element in the human osteocalcin gene," Cell 61:497-504; Yang-Yen, H.F., J.C. Chambard, Y.L. Sun, T. Smeal, T.J. Schmidt, J. Drouin, and M. Karin (1990) "Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction," Cell 62:1205-1215; Yang-Yen, H.F., X.K. Zhang, G. Graupner, M. Tzukerman, B. Sakamoto, M. Karin, and M. Pfahl (1991) "Antagonism between retinoic acid receptors and AP-1: implications for tumor promotion and inflammation," New Biol. 3:1206-1219; Zhang, X.K., K.N. Wills, M. Husmann, T. Hermann, and M. Pfahl (1991) "Novel pathway for thyroid hormone receptor action through interaction with jun and fos oncogene

activities," Mol. Cell Biol. 11:6016-6025). Several reports also suggest a cross-talk between ER and AP-1, however there is no evidence for estrogen dependent inhibition of AP-1 activity (Gaub, M.P., M. Bellard, I. Scheuer, P. Chambon, and P. Sassone-Corsi (1990) "Activation of the ovalbumin gene by the estrogen receptor involves the fos-jun complex," Cell 63:1267-1276; Tzukerman, M., X.K. Zhang, and M. Pfahl (1991) "Inhibition of estrogen receptor activity by the tumor promoter 12-O-tetradecanoylphorbol-13-acetate: a molecular analysis," Mol. Endocrinol. 5:1983-1992). Taken together, it is highly unlikely that AP-1 plays a role in the negative regulation of IL-6 expression by estrogen.

Example 3. Screening agents that differentially affect distinct complexes with the IL-6 promoter

Distinctive complexes with the IL-6 promoter

Treatment of +/-LDA11 cells with IL-1 and TNF $\alpha$  induced binding of NF $\kappa$ B or related proteins to the IL-6 promoter. Since pretreatment with estradiol not only increased the intensity but also the mobility of the complexes, we investigated the binding of +/-LDA11 nuclear extracts to the oligonucleotide covering the putative NF $\kappa$ B site (-82 to -47).

+/-LDA11 cells were pretreated with 10 nM estradiol as indicated. After 24 h the cells were induced with TNF $\alpha$  and IL-1 (1 nM each) for various periods of time. Induction was stopped by cell lysis and nuclear extracts were analyzed by EMSA using the -82 to -47 IL-6 promoter fragment as probe. Fig. 2a shows that extracts from cells treated with IL-1 and TNF $\alpha$  exhibited 3 induced complexes (A,B,C) when compared with extracts from untreated cells.

Over the course of induction (10-120 min) in particular the fastest migrating complex (C) decreased in intensity. Interestingly, estradiol pretreatment reduced the intensity of the slowest migrating complex (A) while strongly increasing the intensity of the fastest band (C). This corresponds to the pattern obtained with the -225 to -52 fragment where complexes seemed to migrate faster with estradiol treatment (Fig. 3).

It is likely that with both fragments analogous complexes were formed; however, only with the shorter oligonucleotide were they completely resolved. Nuclear extract from +/-LDA11 cells treated with estradiol (10 nM) and TNF $\alpha$  and IL-1 (1nM each for 40 min) as indicated were incubated with the -82 to -47 probe in the absence or presence of a 100-fold molar excess of oligonucleotides corresponding to the regions of -82 to -47 and -172 to -131 of the human IL-6 promoter or the vitellogenin ERE. Complexes formed were analyzed by EMSA.

All three induced complexes (A,B,C) were specific since their formation was abolished by inclusion of a 100-fold excess of the unlabeled probe (Fig. 2b, lanes 6 and 10), while the same molar excess of the NF-IL6 oligonucleotide (-172 to -131) or the vitellogenin ERE had no effect (Fig. 2b, lanes 7, 8, 11, and 12).

Fig. 2c shows that when binding of the extracts to the oligonucleotide covering the NF-IL6 site (-172 to -131) was investigated, several complexes were detected. Nuclear extract from +/-LDA11 cells treated with estradiol (10 nM) and TNF $\alpha$  and IL-1 (1nM each for 40 min) as indicated were incubated with the -172 to -131 probe in the absence or presence of a 100-fold molar excess the unlabeled oligonucleotide. Complexes formed were analyzed by EMSA. The arrows in Fig. 2c indicate

the complexes A, B, and C formed upon induction with TNF $\alpha$  and IL-1.

Two of the complexes were specific, since they could be competed with an excess of the unlabeled oligonucleotide (lanes 2, 4, and 6). However, all of the complexes were formed constitutively, independent of cytokine induction or estradiol treatment, suggesting that they were unrelated to the regulation of IL-6 expression by IL-1, TNF $\alpha$ , and estrogen.

#### Screening for compounds that affect the formation of distinct complexes

In more detailed studies we analyzed the effects of other compounds on the formation of complexes A, B, and C (Fig. 3). +/-LDA11 cells were pretreated with cycloheximide (CHX) or the kinase inhibitor H7 for 5 min or with estradiol (10 nM) and/or ICI 164,384 (100 nM) for 24 h or 60 min before induction with TNF $\alpha$  and IL-1 (1nM each for 30 min). Treatment was stopped by cell lysis and nuclear extracts were analyzed by EMSA using the -82 to -47 fragment as probe. Arrows indicate the induced complexes A, B, and C.

Pretreatment of the cells with the protein synthesis inhibitor cycloheximide before addition of cytokines did not interfere with complex formation (lane 8). This is consistent with the fast induction of binding and has been shown before for the activation of NF $\kappa$ B (Henkel, T., T. Machleidt, I. Alkalay, M. Kronke, Y. Ben-Neriah, and P.A. Baeuerle (1993) "Rapid proteolysis of I kappa B-alpha is necessary for activation of transcription factor NF-kappa B," Nature 365:182-185; Sen, R. and D. Baltimore (1986) "Inducibility of kappa immunoglobulin enhancer-binding protein Nf-kappa B by a posttranslational mechanism," Cell 47:921-928). It has been reported that cycloheximide treatment activates NF $\kappa$ B binding (Sen, R. and D. Baltimore (1986)



"Inducibility of kappa immunoglobulin enhancer-binding protein NF-kappa B by a posttranslational mechanism," Cell 47:921-928; Zhang, Y.H., J.X. Lin, and J. Vilcek (1990) "Interleukin-6 induction by tumor necrosis factor and interleukin-1 in human fibroblasts involves activation of a nuclear factor binding to a kappa B-like sequence," Mol. Cell Biol. 10:3818-3823). However, in +/+LDA11 cells we did not observe any induction by cycloheximide (lane 7). Additionally, pretreatment with H-7, a potent protein kinase C (PKC) inhibitor (Kawamoto, S. and H. Hidaka (1984) "1-(5-Isoquinolinesulfonyl)-2-methylpiperazine (H-7) is a selective inhibitor of protein kinase C in rabbit platelets," Biochem. Biophys. Res. Commun. 125:258-264), did not interfere with induction of complex formation by IL-1 and TNF $\alpha$  (lane 10). This suggests that induction of complexes A, B, and C is mediated through a PKC-independent pathway and is consistent with the finding that IL-1 and TNF $\alpha$  activate NF $\kappa$ B and induce IL-6 independently of PKC (11,55,99).

Activation of NF $\kappa$ B by phorbol esters is probably mediated by PKC (Baeuerle, P.A. and D. Baltimore (1988) "I kappa B: a specific inhibitor of the NF-kappa B transcription factor," Science 242:540-546). However, treatment of +/+LDA11 cells with 12-O-tetradecanoylphorbol-13-acetate (TPA) did not significantly induce production of IL-6, nor did it induce complex formation with the -82 to -47 fragment.

The estradiol effect, i.e. decreasing complex A and increasing complex C (Fig. 3, compare lanes 2 and 3), was not seen with a short estradiol pretreatment (60 min) before induction (lane 6). In addition, the pure anti-estrogen ICI 164,384 (Wakeling, A.E. and J. Bowler (1988) "Biology and mode of action of pure antioestrogens," J. Steroid Biochem. 30:141-147) did not affect the complex pattern (lane 4). However, when ICI

164,384 was added in combination with estradiol it prevented the effect mediated by the estrogen (lane 5).

Since ICI 164,384 acts as an antagonist via binding to the ER, these results further support the hypothesis that the effects of estradiol on the induced complexes are receptor mediated. However, the mechanism of estrogen action is probably indirect, as indicated by the lack of response to short term estradiol treatment.

#### Screening for agents that affect binding characteristics of the proteins in distinct complexes

To investigate the binding characteristics of the proteins in complexes A, B, and C with the NF $\kappa$ B oligonucleotide, methylation interference experiments were carried out (Fig. 4). Nuclear +/-LDA11 extracts from cells induced with TNF $\alpha$  and IL-1 (1nM each) were incubated with -82 to -47 probe that had been labeled either on the upper or the lower strand and subjected to limited DMS-methylation. After preparative EMSA, DNA from complexes A, B, and C and from the unretarded probe (F) was isolated, cleaved with piperidine, and electrophoresed on a 12% denaturing gel. The sequence corresponding to the NF $\kappa$ B consensus site is shown boxed, a cryptic AP-1 site is shaded.

On both strands N-7-methylation of the guanine bases within the NF $\kappa$ B site (-73 to -63, boxed) interfered with complex formation, while methylation of guanines flanking the consensus site had no observable effect. The interference pattern for all three complexes (A,B,C) was identical.

The observation that methylation of guanines just outside of the NF $\kappa$ B site (-75, -60, -58) did not affect the formation of even the largest complex (A) suggests that in all three complexes DNA contacts are made within the same core region. In addition, the lack of DNA binding interference with methylation of guanines -60, -

58, and -56 strongly argues against any cytokine induced binding of factors to the nonconsensus (TGAGTCT, shaded) AP-1 site (Tanabe, O., S. Akira, T. Kamiya, G.G. Wong, T. Hirano, and T. Kishimoto (1988) "Genomic structure of the murine IL-6 gene. High degree conservation of potential regulatory sequences between mouse and human," J. Immunol. 141:3875-3881) in this region (-61 to -55).

Our studies show that estrogen affects the formation of complexes with the IL-6 promoter that involve NF $\kappa$ B p50 and p65 or very closely related proteins. Treatment of +/+LDA11 cells with IL-1 and TNF $\alpha$  specifically induced the formation of at least three distinct complexes with the NF $\kappa$ B consensus site in the IL-6 promoter. Although of various size, in all three complexes the DNA contacts are restricted to the core sequence of the NF $\kappa$ B site. The corresponding core sequence of other NF $\kappa$ B elements is protected by p50 and p65 (Baldwin, A.S., Jr. and P.A. Sharp (1988) "Two transcription factors, NF-kappa B and H2TF1, interact with a single regulatory sequence in the class I major histocompatibility complex promoter," Proc. Natl. Acad. Sci. USA 85:723-727; Kieran, M., V. Blank, F. Logeat, J. Vandekerckhove, F. Lottspeich, O. Le Bail, M.B. Urban, P. Kourilsky, P.A. Baeuerle, and A. Israel (1990) "The DNA binding subunit of NF-kappa B is identical to factor KBF1 and homologous to the rel oncogene product," Cell 62:1007-1018; Sen, R. and D. Baltimore (1986) "Multiple nuclear factors interact with the immunoglobulin enhancer sequences," Cell 46:705-716), both of which have been shown to directly interact with DNA (Nolan, G. P., S. Ghosh, H.C. Liou, P. Tempst, and D. Baltimore (1991) "DNA binding and I kappa B inhibition of the cloned p65 subunit of NF-kappa B, a rel-related polypeptide," Cell 64:961-969; Urban, M.B., R. Schreck, and P.A. Baeuerle (1991) "NF-kappa B contacts DNA by a heterodimer of the p50 and p65 subunit," EMBO J.

10:1817-1825). C-rel homodimers and heterodimers with p50 have been shown to bind the NF $\kappa$ B site in the IL-6 promoter (Nakayama, K., H. Shimizu, K. Mitomo, T. Watanabe, S. Okamoto, and K. Yamamoto (1992) "A lymphoid cell-specific nuclear factor containing c-Rel-like proteins preferentially interacts with interleukin-6 kappa B-related motifs whose activities are repressed in lymphoid cells," Mol. Cell Biol. 12:1736-1746).

In addition, this study showed that in lymphoid cells c-rel or an immunologically related factor is a component of a larger complex that binds the NF $\kappa$ B site in the IL-6 promoter and functions as a constitutive repressor. In +/+LDA11 cells, we could not detect any c-rel specific binding activity. A number of other NF $\kappa$ B unrelated proteins have been shown to bind to NF $\kappa$ B consensus sites. Those include  $\alpha$ A-CRYBP1 (Nakamura, T., D.M. Donovan, K. Hamada, C.M. Sax, B. Norman, J.R. Flanagan, K. Ozato, H. Westphal, and J. Piatigorsky (1990) "Regulation of the mouse alpha A-crystallin gene: isolation of a cDNA encoding a protein that binds to a cis sequence motif shared with the major histocompatibility complex class I gene and other genes," Mol. Cell Biol. 10:3700-3708), MBP-1/PRDII-BFI (Baldwin, A.S., Jr., K.P. LeClair, H. Singh, and P.A. Sharp (1990) "A large protein containing zinc finger domains binds to related sequence elements in the enhancers of the class I major histocompatibility complex and kappa immunoglobulin genes," Mol. Cell Biol. 10:1406-1414; Fan, C.M. and T. Maniatis (1990) "A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence," Genes Dev. 4:29-42), and AGIE-BP1 (Ron, D., A.R. Brasier, and J.F. Habener (1991) "Angiotensinogen gene-inducible enhancer-binding protein 1, a member of a new family of large nuclear proteins that recognize nuclear

factor kappa B-binding sites through a zinc finger motif," Mol. Cell Biol. 11:2887-2895).

It has been shown that C/EBP-like proteins attenuate NF $\kappa$ B mediated transactivation of the angiotensinogen gene acute-phase response element (Brasier, A.R., D. Ron, J.E. Tate, and J.F. Habener (1990) "A family of constitutive C/EBP-like DNA binding proteins attenuate the IL-1 alpha induced, NF kappa B mediated trans-activation of the angiotensinogen gene acute-phase response element," EMBO J. 9:3933-3944). Currently, we cannot exclude that those proteins or others are part of the observed complexes or are involved in the inhibition of IL-6 expression by estrogen. A recent study suggested that in uterine cells, estradiol induced complex formation with an NF $\kappa$ B element (Shyamala, G. and M.C. Guiot (1992) "Activation of kappa B-specific proteins by estradiol," Proc. Natl. Acad. Sci. USA 89:10628-10632). The induced complex did not contain p50 or p65 and therefore may represent other factors.

#### Example 4. Screening for agents that affects the binding of p65 to IL-6 promoter

##### Analyzing composition of the complexes formed with the NF $\kappa$ B site

As with the larger promoter fragment, we analyzed the nature of the complexes formed with the NF $\kappa$ B oligonucleotide (-82 to -47) by antibody shift experiments (Fig. 5a). Nuclear extract from +/-LDA11 cells treated with estradiol (10 nM) and TNF $\alpha$  and IL-1 (1nM each for 10 min) as indicated were incubated with the -82 to -47 probe in the absence or presence of various antibodies. Complexes formed were analyzed by EMSA.

We observed strong effects on the induced complexes when anti-p50 or anti-p65 (lanes 17-20) were included in

the binding reactions. Interestingly, anti-p50 specifically abolished the formation of complexes B and C, seemed to leave complex A unaffected, and caused the appearance of a single supershifted band (S1). Anti-p65, however, inhibited the formation of all three induced complexes and produced two supershifted bands (S1, S2). This suggested that p65 or an immunologically closely related protein is part of all three induced complexes, while p50 or a related protein is only present in complexes B and C.

It has been reported that recombinantly expressed c-rel binds to the NF $\kappa$ B site in the IL-6 promoter as heterodimer with p50 and, with particular high affinity, as homodimer (Nakayama, K., H. Shimizu, K. Mitomo, T. Watanabe, S. Okamoto, and K. Yamamoto (1992) "A lymphoid cell-specific nuclear factor containing c-Rel-like proteins preferentially interacts with interleukin-6 kappa B-related motifs whose activities are repressed in lymphoid cells," Mol. Cell Biol. 12:1736-1746). When anti-c-rel was included in the binding reactions with the +/+LDA11 extracts the antibody did not inhibit any of the induced complexes. On longer exposures a weak supershifted complex was detectable. This complex migrated at the same position as the supershift observed with anti-p50 suggesting that it did not contain the larger c-rel protein. Since the peptide used to raise the anti-c-rel antibody has a 56% homology to the analogous p50 sequence (Ghosh, S., A.M. Gifford, L.R. Riviere, P. Tempst, G.P. Nolan, and D. Baltimore (1990) "Cloning of the p50 DNA binding subunit of NF-kappa B: homology to rel and dorsal," Cell 62:1019-1029; Inoue, J., L.D. Kerr, L.J. Ransone, E. Bengal, T. Hunter, and I.M. Verma (1991) "c-rel activates but v-rel suppresses transcription from kappa B sites," Proc. Natl. Acad. Sci. USA 88:3715-3719), it is likely that this weak band is the result of a cross-reactivity and unrelated to c-

rel. As with the larger promoter fragment, anti-c-jun did not affect complex formation (lanes 13 and 14). Results shown depict the 10 min induction time point and are essentially identical with longer cytokine treatments.

None of the antibodies tested had a marked effect on complex formation with the uninduced extracts. Only anti-p50 produced a very weak supershifted complex migrating at the S1 position as observed with the induced extracts (only visible on longer exposures). This weak binding activity, only detectable when supershifted, could either result from a cytosolic contamination or represent basal activation under the culture conditions.

In additional EMSA experiments we included a purified preparation of recombinant p50 as well as the yeast expressed ER (Fig 5b). Nuclear +/-LDA11 extracts from cells pretreated with estradiol (10 nM) and induced with TNF $\alpha$  and IL-1 (1nM each for 30 min) as indicated as well as purified human p50 protein and yeast extract containing recombinantly expressed human ER were incubated with -82 to -47 probe in the absence or presence of various antibodies. Complexes formed were analyzed by EMSA. Arrows indicate complexes A, B, and C induced by cytokine treatment and the complexes S1 and S2 resulting from the antibody supershifts.

As expected, neither did yeast expressed ER bind to the NF $\kappa$ B -82 to -47 fragment (lanes 5, 9, 13, and 17) nor was ER involved in the formation of the induced complexes as indicated by the lack of any anti-ER antibody effect (lanes 14-17). Purified p50 bound to this fragment and was specifically supershifted by anti-p50 but not by anti-p65 (compare lanes 4, 8, and 12).

Surprisingly, the complex formed with purified recombinant p50 migrated slower than complexes B and C. Using lower concentrations of purified p50 did not

affect the migration, suggesting that the band represented p50 homodimers and not higher order complexes (Duckett, C.S., N.D. Perkins, T.F. Kowalik, R.M. Schmid, E.S. Huang, A.S. Baldwin, Jr., and G.J. Nabel (1993) "Dimerization of NF-KB2 with RelA(p65) regulates DNA binding, transcriptional activation, and inhibition by an I kappa B-alpha (MAD-3)," Mol. Cell Biol. 13:1315-1322). However, the antibody shift experiments suggested that both NF-kB proteins, p50 and p65, are part of complexes B and C (lanes 6, 7, 11, and 12) and consequently both complexes should migrate slower than p50 homodimers (Urban, M.B., R. Schreck, and P.A. Baeuerle (1991) "NF-kappa B contacts DNA by a heterodimer of the p50 and p65 subunit," EMBO J. 10:1817-1825).

It is possible that the proteins in complexes B and C are only immunologically related to p50 and p65 but actually of smaller size. Speculations that the p50 homologue p49 is part of the induced complexes and is responsible for the faster migration could not be confirmed. Although purified p49 bound the -82 to -47 IL-6 fragment and strongly cross-reacted with the anti-p50 antibody, the complex formed migrated even more slowly than the p50 complex. This corresponds to results obtained with other NF-kB binding sites (Duckett, C.S., N.D. Perkins, T.F. Kowalik, R.M. Schmid, E.S. Huang, A.S. Baldwin, Jr., and G.J. Nabel (1993) "Dimerization of NF-KB2 with RelA(p65) regulates DNA binding, transcriptional activation, and inhibition by an I kappa B-alpha (MAD-3)," Mol. Cell Biol. 13:1315-1322).

However, the finding that the anti-p50 antibody strongly cross-reacts with p49 indicates that the antibodies used may detect other NF-kB related proteins in the complexes. Alternatively, the migration of complexes B and C could be higher than the migration



observed with recombinant p50 due to conformational differences resulting from post-translational modification. This would correlate with the observation that the inclusion of anti-p50, abolishing complexes B and C, produced a supershifted complex (S1) migrating more slowly than the supershifted complex obtained with recombinant p50 (Fig. 5b, lanes 6-8).

A closer inspection of the band shift results indicated that anti-p50 also affected complex A. As discussed before, treatment with estradiol not only increased the intensity of complex C but also decreased the intensity of complex A (lanes 2 and 3). We consistently observed that inclusion of anti-p50 had a very similar effect: the antibody specifically decreased the intensity of complex A formed with the extracts from cells not treated with estradiol, resulting in equal intensity of this band using extracts from estradiol treated or untreated cells (Fig. 5b, compare lanes 2 and 7). These results suggest that band A induced by IL-1 and TNF $\alpha$  in the absence of estradiol is composed of two different unresolved complexes, one (A1) that is also induced in the presence of estradiol and does not contain p50, and another complex (A2) containing p50 (or an immunologically related protein).

Treatment with estradiol may increase the intensity of complex C at the expense of complex A2. If complex A2 represents a transcriptionally more active state, this could explain the inhibitory effect of estradiol on IL-6 expression. Recently it has been shown that the TATA binding protein (TBP or TFIIDt) directly interacts with NF $\kappa$ B (Kerr, L.D., L.J. Ransone, P. Wamsley, M.J. Schmitt, T.G. Boyer, Q. Zhou, A.J. Berk, and I.M. Verma (1993) "Association between proto-oncoprotein Rel and TATA-binding protein mediates transcriptional activation by NF-kappa B," *Nature* 365:412-419). Therefore, we were interested if TBP was involved in the formation of the

induced complexes. However, using an anti-TBP antibody we could not detect any participation of TBP in complexes A, B, or C.

Our antibody gel shift experiments suggested that p65 is a component of all three observed complexes. This particular protein is the NF $\kappa$ B component containing the transactivation domain (Schmitz, M.L. and P.A. Baeuerle (1991) "The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B," EMBO J. 10:3805-3817). Within the different complexes the transactivation function may be differentially active. The antibody shift experiments suggest that estradiol diminishes the A2 complex while increasing complex C. The slow migrating A2 complex may contain other factor(s) involved in the transactivation process. The TATA-binding protein TBP, part of the TFIID complex has been reported to interact strongly with c-rel and p65, but not with p50 or p49 (Kerr, L.D., L.J. Ransone, P. Wamsley, M.J. Schmitt, T.G. Boyer, Q. Zhou, A.J. Berk, and I.M. Verma (1993) "Association between proto-oncoprotein Rel and TATA-binding protein mediates transcriptional activation by NF-kappa B," Nature 365:412-419). However, using a TBP-specific antibody we could not detect TBP as part of any of the complexes formed with the NF $\kappa$ B site in the IL-6 promoter.

#### Example 5. Efficacy-testing of Putative Cytokine Modulators

Methods for testing the efficacy of putative cytokine modulators are provided. Each candidate compound is tested for its efficacy in modulating cytokine expression in cell lines, in animal models, and in controlled clinical studies using methods known to those skilled in the art and approved by the Food and Drug Administration, such as, but not limited to, those

promulgated in The Federal Register 47 (no. 56): 12558-12564, March 23, 1982.

Example 6. Toxicity-testing of Putative Cytokine  
Modulators

Methods are provided for determining whether an agent active in any of the methods listed above has little or no effect on healthy cells. Such agents are then formulated in a pharmaceutically acceptable buffer or in buffers useful for standard animal tests.

By "pharmaceutically acceptable buffer" is meant any buffer which can be used in a pharmaceutical composition prepared for storage and subsequent administration, which comprise a pharmaceutically effective amount of an agent as described herein in a pharmaceutically acceptable carrier or diluent. Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co. (A.R. Gennaro edit. 1985). Preservatives, stabilizers, dyes and even flavoring agents may be provided in the pharmaceutical composition. For example, sodium benzoate, sorbic acid and esters of p-hydroxybenzoic acid may be added as preservatives. Id. at 1449. In addition, antioxidants and suspending agents may be used. Id.

A. Additional screens for Toxicity: Method 1

Agents identified as having cytokine modulating activity are assessed for toxicity to cultured human cells. This assessment is based on the ability of living cells to reduce 2,3,-bis[2-methoxy-4-nitro-5-sulphonylphenyl]-5-[(phenylamino)carbonyl]-2H-tetrazolium hydroxide] otherwise referred to as XTT (Paull et al., J. Heterocycl. Chem. 25:763-767 (1987); Weislow et al., (1989), J. Natl. Canc. Inst. 81:577). Viable mammalian cells are capable of reductive cleavage of an N-N bond

in the tetrazole ring of XTT to form XTT formazan. Dead cells or cells with impaired energy metabolism are incapable of this cleavage reaction. The extent of the cleavage is directly proportional to the number of living cells tested. Cells from a human cell line such as HeLa cells are seeded at  $10^3$  per well in 0.1 ml of cell culture medium (Dulbecco's modified minimal essential medium supplemented with 10% fetal calf serum) in the wells of a 96 well microtiter plate. Cells are allowed to adhere to the plate by culture at 37° C in an atmosphere of 95% air, 5% CO<sub>2</sub>. After overnight culture, solutions of test substances are added in duplicate to wells at concentrations that represent eight half-decade log dilutions. In parallel, the solvent used to dissolve the test substance is added in duplicate to other wells. The culture of the cells is continued for a period of time, typically 24 hours. At the end of that time, a solution of XTT and a coupler (methylphenazonium sulfate) is added to each of the test wells and the incubation is continued for an additional 4 hours before the optical density in each of the wells is determined at 450 nm in an automated plate reader. Substances that kill mammalian cells, or impair their energy metabolism, or slow their growth are detected by a reduction in the optical density at 450 nm in a well as compared to a well which received no test substance.

B. Additional screens for Toxicity: Method 2

Cytokine modulators are tested for cytotoxic effects on cultured human cell lines using incorporation of <sup>35</sup>S methionine into protein as an indicator of cell viability. HeLa cells are grown in 96 well plates in Dulbecco's minimal essential medium supplemented with 10% fetal calf serum and 50 µg/ml penicillin and streptomycin. Cells are initially seeded at  $10^3$  cells/well, 0.1 ml/well. Cells are grown for 48 hrs without exposure to the cytokine modulator, then medium

is removed and varying dilutions of the cytokine modulator prepared in complete medium are added to each well, with control wells receiving no cytokine modulator. Cells are incubated for an additional 48-72 hrs. Medium is changed every 24 hrs and replaced with fresh medium containing the same concentration of the cytokine modulators. Medium is then removed and replaced with complete medium without antifungal. Cells are incubated for 24 hr in the absence of cytokine modulator, then viability is estimated by the incorporation of  $^{35}\text{S}$  into protein. Medium is removed, replaced with complete medium without methionine, and incubated for 30 min. Medium is again removed, and replaced with complete medium without methionine but containing  $0.1 \mu\text{Ci/ml}$   $^{35}\text{S}$  methionine. Cells are incubated for 3 hrs. Wells are washed 3 times in PBS, then cells are permeabilized by adding 100% methanol for 10 min. Ice cold 10% trichloroacetic acid (TCA) is added to fill wells; plates are incubated on ice for 5 min. This TCA wash is repeated two more times. Wells are again washed in methanol, then air dried.  $50 \mu\text{l}$  of scintillation cocktail are added to each well and dried onto the wells by centrifugation. Plates are used to expose X ray film. Densitometer scanning of the autoradiogram, including wells without antifungal, is used to determine the dosage at which 50% of cells are not viable ( $\text{ID}_{50}$ ) (Culture of Animal Cells. A manual of basic technique. (1987). R. Ian Freshney. John Wiley & Sons, Inc., New York).

#### Example 7. Administration of Cytokine Modulators

The invention features novel cytokine modulators discovered by the methods described above. It also includes novel pharmaceutical compositions which include cytokine modulators discovered as described above formulated in pharmaceutically acceptable formulations.

Furthermore, the invention features a method for treating a subject inflicted with a pathological condition affected by the level of a cytokine by administering to that subject a therapeutically effective amount of a cytokine modulator. Such administration can be by any method known to those skilled in the art, for example, by topical application or by systemic administration.

By "therapeutically effective amount" is meant an amount that relieves (to some extent) one or more symptoms of the disease or condition in the patient. Additionally, by "therapeutically effective amount" is meant an amount that returns to normal, either partially or completely, physiological or biochemical parameters associated with or causative of a mycotic disease or condition. Generally, it is an amount between about 1 nmole and 1  $\mu$ mole of the molecule, dependent on its  $EC_{50}$  and on the age, size, and disease associated with the patient.

Other embodiments of this invention are disclosed in the following claims.

WHAT IS CLAIMED IS

1. Method for screening for a therapeutic agent for treatment of a pathological condition affected by the level of a cytokine, comprising the steps of:

contacting a potential therapeutic agent with a  
5 system comprising an intracellular receptor, a promoter or a portion of said promoter with a rel site, and a protein that binds to said rel site on said promoter;  
measuring the binding of said protein to said rel site on said promoter; wherein a reduction in the  
10 binding of said protein to the rel site on the promoter compared to the binding of said protein in the absence of said agent is an indication that said agent is potentially useful for treatment of said condition.

2. The method of claim 1, wherein said protein is  
15 a rel-like protein.

3. The method of claim 2, wherein said rel-like protein is NF $\kappa$ B.

4. The method of claim 1, wherein said system further comprises a ligand for said intracellular  
20 receptor.

5. The method of claim 1, wherein said condition is osteoporosis.

6. The method of claim 1, wherein said condition is rheumatoid arthritis.

25 7. The method of claim 1, wherein said condition is inflammation.

8. The method of claim 1, wherein said condition is psoriasis.

9. The method of claim 1, wherein said condition is Kaposi's sarcoma.

---

10. The method of claim 1, wherein said condition is septic shock.

5 11. The method of claim 1, wherein said condition is multiple myeloma.

12. The method of claim 1, wherein said intracellular receptor is a steroid receptor.

10 13. The method of claim 1, wherein said intracellular receptor is an estrogen receptor.

14. The method of claim 1, wherein said intracellular receptor is selected from the group consisting of retinoid acid receptors, retinoid X receptors, glucocorticoid receptor, progesterone  
15 receptors, androgen receptor, thyroid hormone receptors, and vitamin D receptor.

15. The method of claim 1, wherein said measuring comprises determining the expression level of a cytokine or an acute phase protein.

20 16. The method of claim 1, wherein said measuring comprises determining the expression level of a reporter gene linked to said promoter.

17. The method of claim 1, wherein said system further comprises an effector of said promoter.

25 18. The method of claim 17, wherein said effector is selected from the group consisting of tumor necrosis factor, interleukin-1, viruses, endotoxin, phorbol



esters, epidermal growth factor, leukemia inhibitor factor and cAMP agonists.

---

19. The method of claim 1, wherein said cytokine is interleukin 6.

5        20. The method of claim 1, wherein said cytokine is interleukin 8.

21. The method of claim 1, wherein said system is a cell.

22. The method of claim 1, wherein said system is  
10 an extract of a cell.

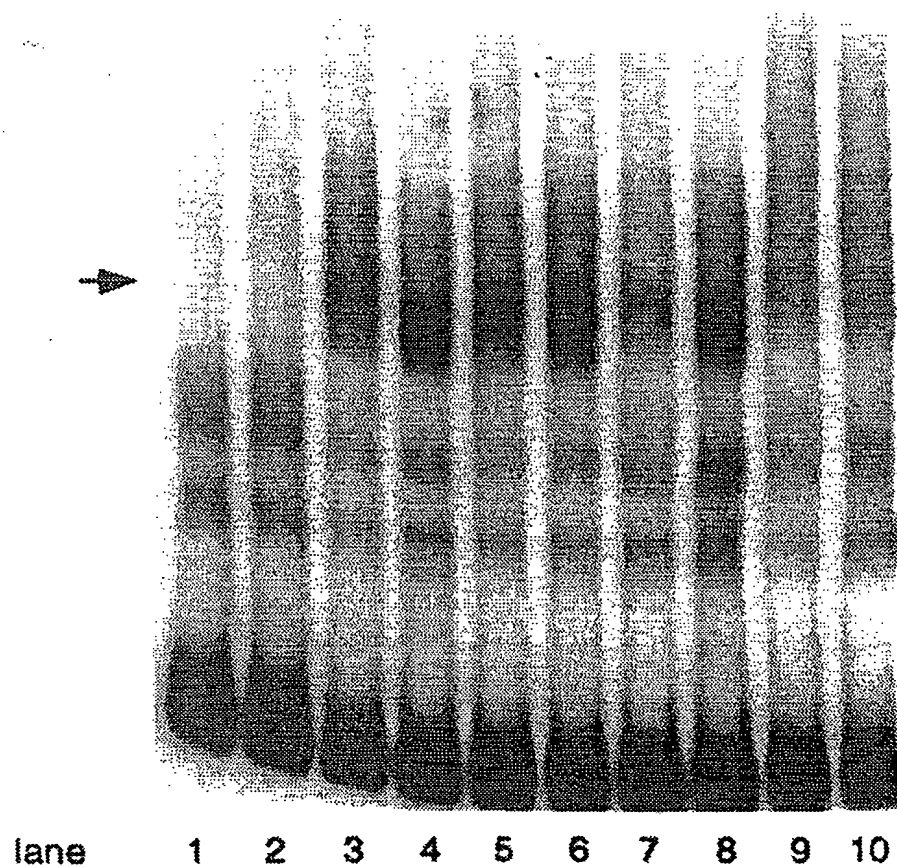
23. The method of claim 21, wherein said intracellular receptor is expressed from a transfected vector.

24. The method of claim 21, wherein said promoter  
15 or said portion of said promoter is transfected into said cell.

I/II

FIG. 1a.

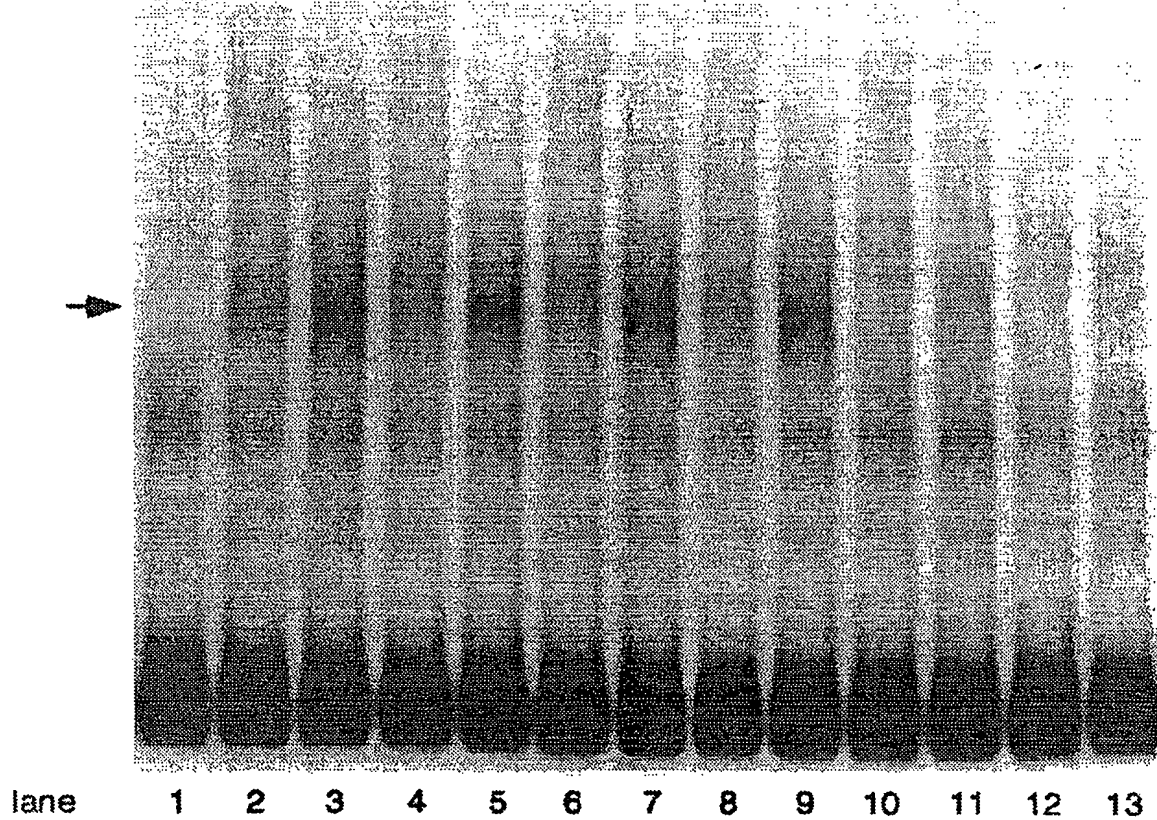
TNF/IL-1	none	10 min	20 min	40 min	120 min
17 $\beta$ -E <sub>2</sub>	- +	- +	- +	- +	- +



2/11

FIG. 1b.

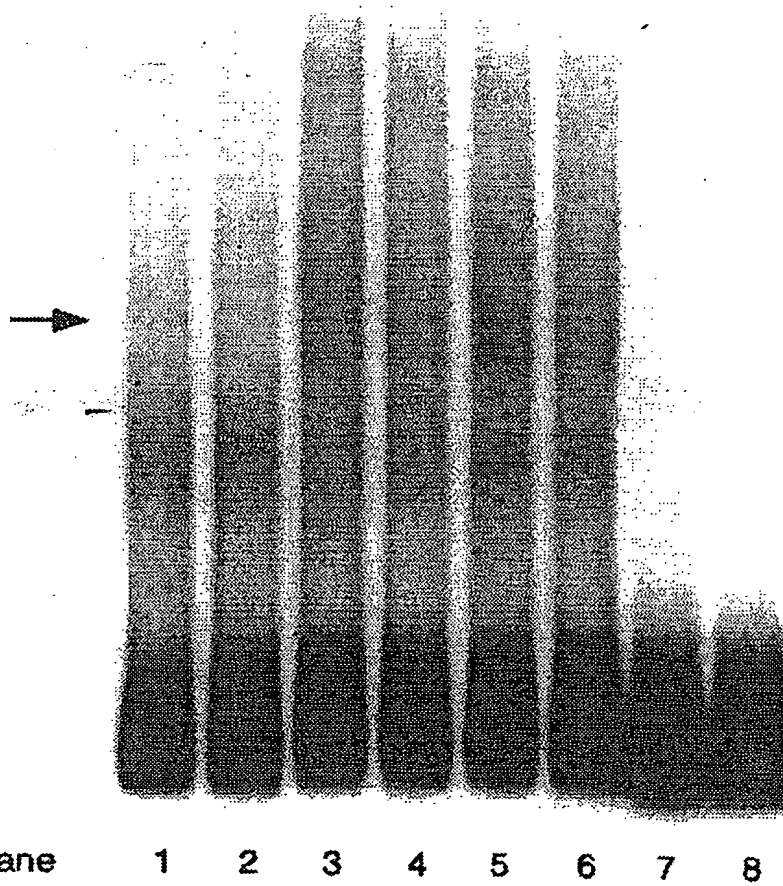
Induction	none		TNF/IL-1									
Antibody	$\alpha$ p65	none		$\alpha$ c-jun	$\alpha$ NF-IL6	$\alpha$ c-rel	$\alpha$ p50	$\alpha$ p65				
17 $\beta$ -E <sub>2</sub>	+	-	+	-	+	-	+	-	+	-	+	-



3/11

FIG. 1c.

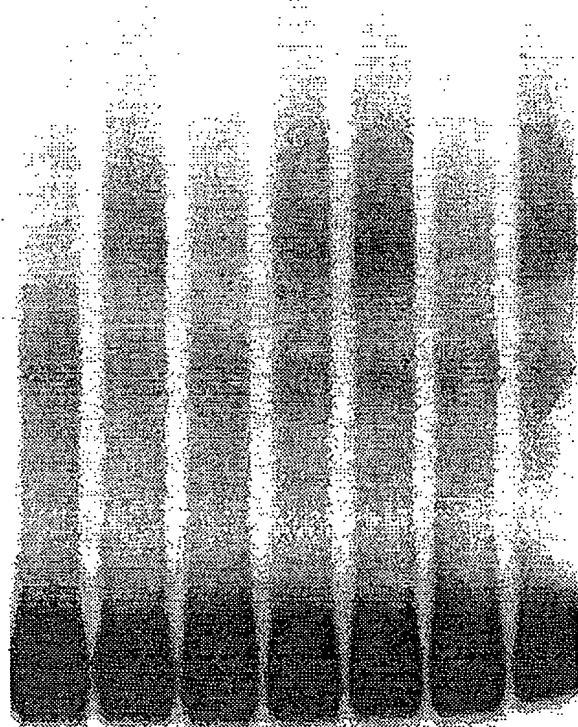
Extract/ protein	+/+LDA11				ER	
TNF/IL-1	-		+		-	
17 $\beta$ -E <sub>2</sub>	+		-		+	
anti-ER	-	+	-	+	-	+



4/11

*FIG. 1d.*

Induction	none	TNF/IL-1					
17 $\beta$ -E <sub>2</sub>	+	-				+	
Competitor	none	none	82-47	172-131	none	82-47	172-131



lane      1      2      3      4      5      6      7

5/11

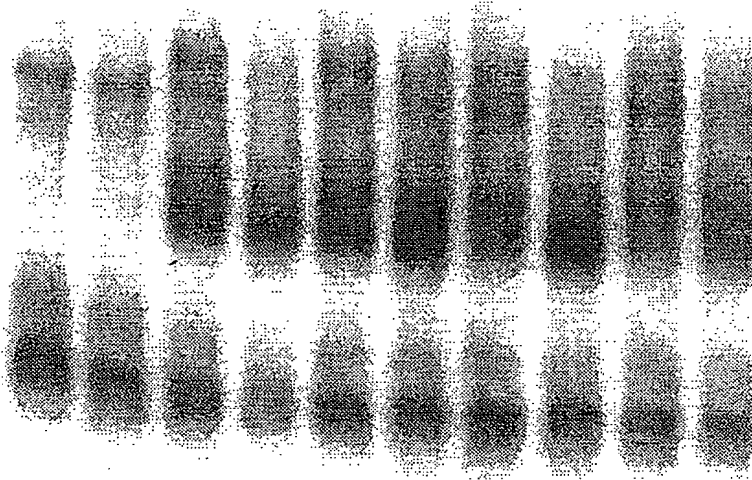
FIG. 2a.

TNF/IL-1	none	10 min	20 min	40 min	120 min
17 $\beta$ -E <sub>2</sub>	- +	- +	- +	- +	- +

A →

B →

C →



lane

1 2 3 4 5 6 7 8 9 10

6/11

FIG. 2b.

Induction 17 $\beta$ -E <sub>2</sub>	none				TNF/IL-1							
	+				-				+			
Competitor	none	82-47	172-131	ERE	none	82-47	172-131	ERE	none	82-47	172-131	ERE

A →

B →

C →

lane

1

2

3

4

5

6

7

8

9

10

11

12

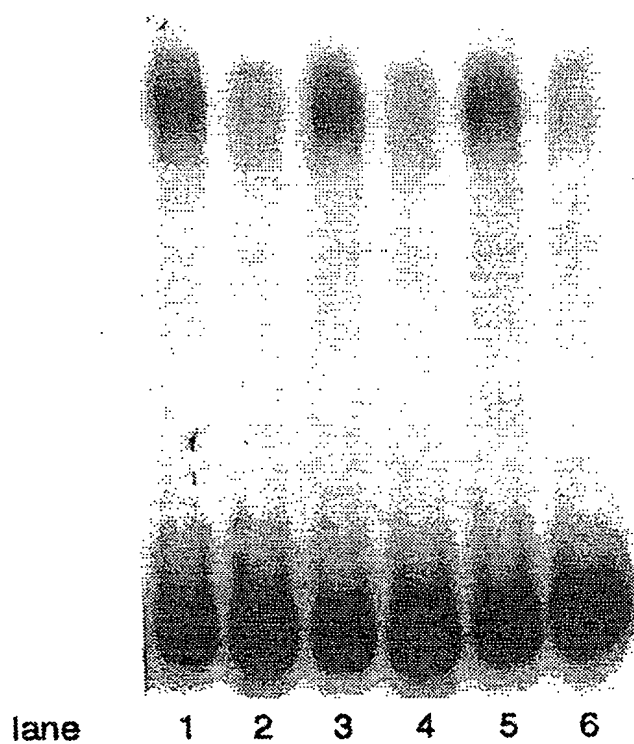
SUBSTITUTE SHEET (RULE 26)

7/11

FIG. 2c.

1/11

Induction 17 $\beta$ -E <sub>2</sub>	none		TNF/IL-1			
	+		-		+	
Competitor	none		172-131		172-131	
	none	172-131	none	172-131	none	172-131





8/11

FIG. 3.

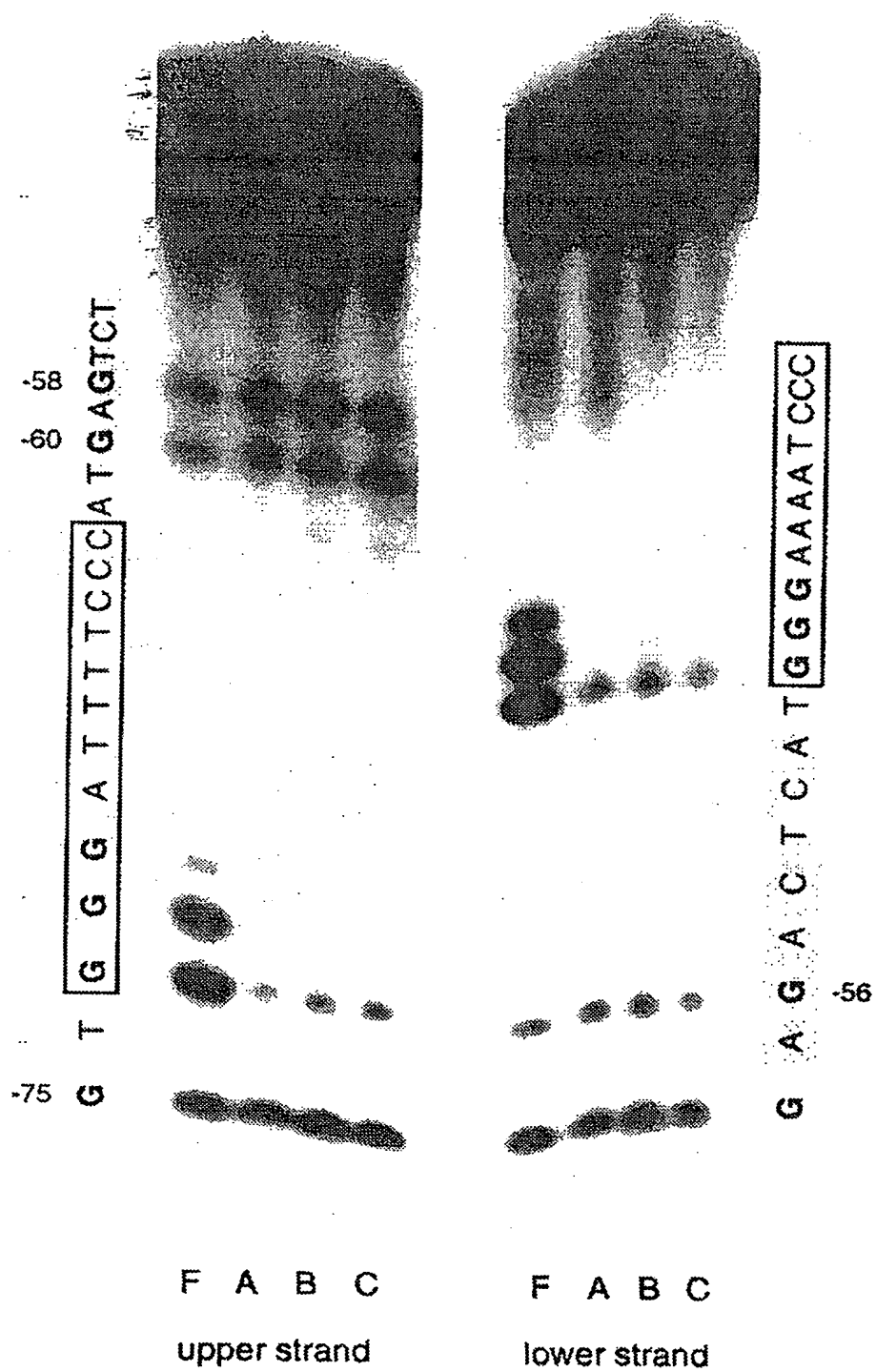
Induction	TNF/IL-1						none	TNF/IL-1		none	TNF/IL-1	
	none	none	E <sub>2</sub>	ICI	E <sub>2</sub> + ICI	60 min E <sub>2</sub>	CHX	CHX	H7	H7	H7	H7
Pretreatment	none	none	E <sub>2</sub>	ICI	E <sub>2</sub> + ICI	60 min E <sub>2</sub>	CHX	CHX	H7	H7	H7	H7

A →  
B →  
C →

lane 1 2 3 4 5 6 7 8 9 10

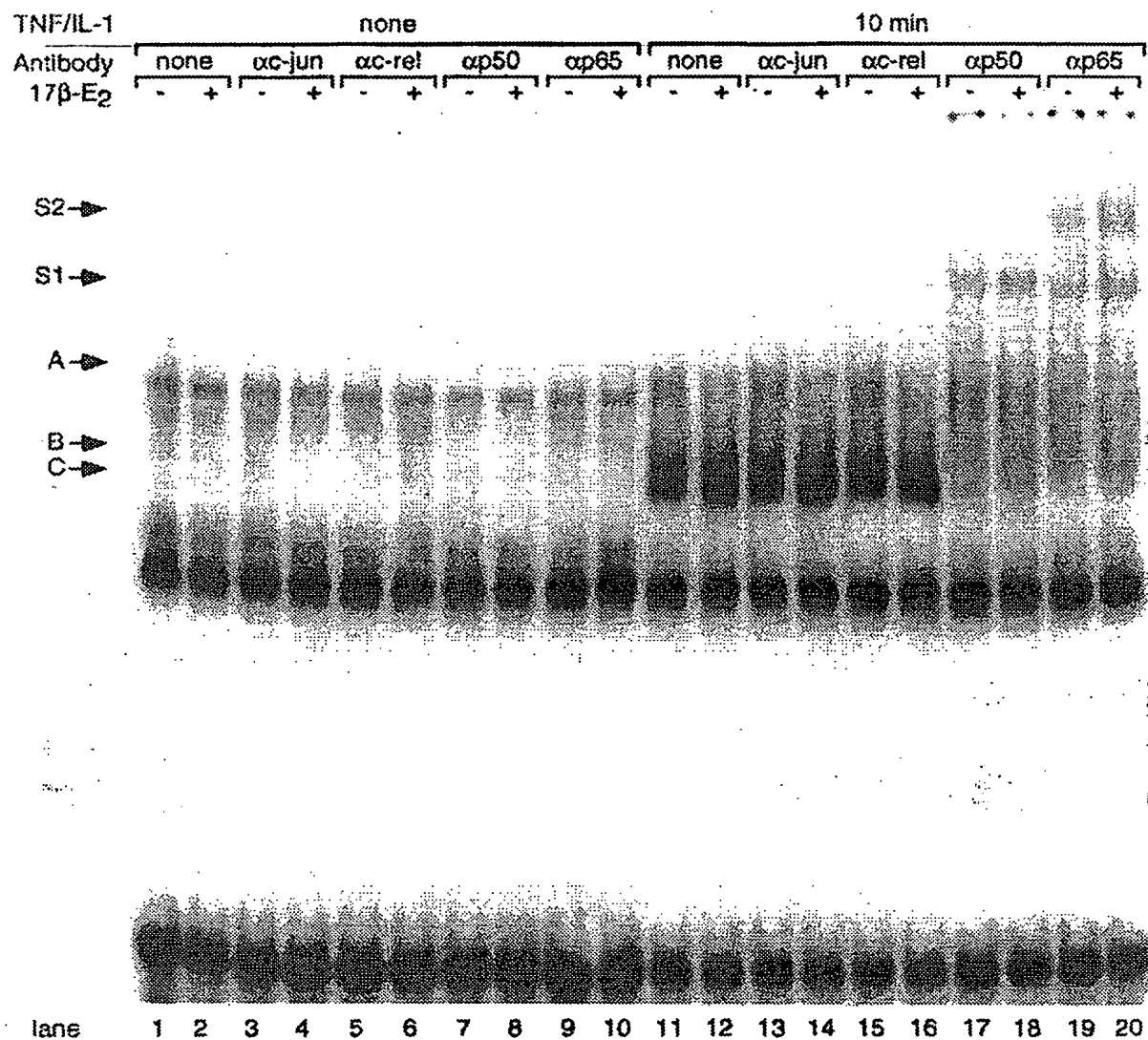
9/11

FIG. 4.



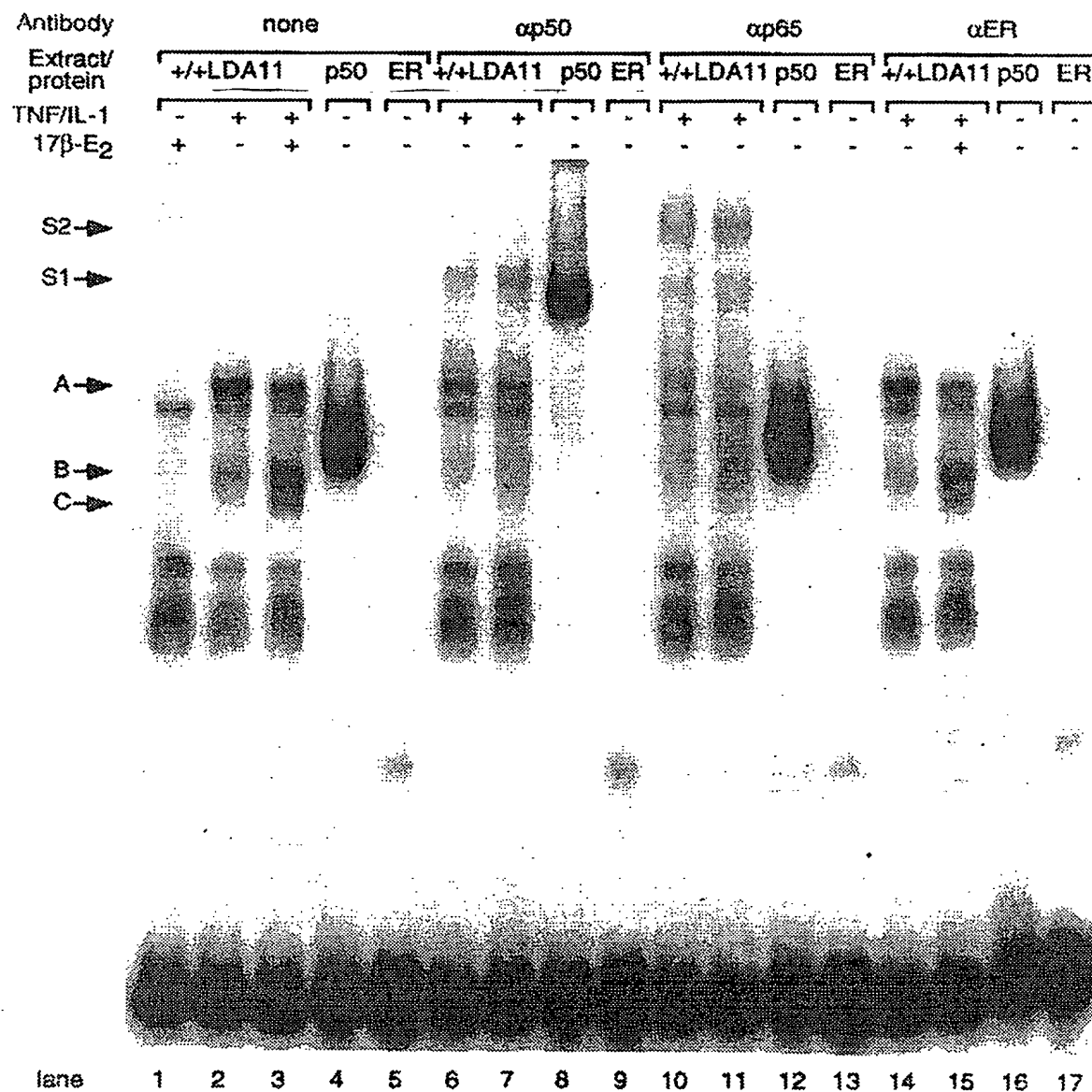
10/11

FIG. 5a.



II/II

FIG. 5b.



## INTERNATIONAL SEARCH REPORT

Internal Application No

PCT/US 95/06524

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 G01N33/50 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO,A,92 07072 (LA JOLLA CANCER RESEARCH FOUNDATION.THE REGENTS OF THE U. CALIFORNIA) 30 April 1992 cited in the application ---	
A	THE EMBO JOURNAL, vol. 9, no. 6, 1990 pages 1897-1906, S, AKIRA ET AL. 'A nuclear factor for IL-6 expression is a member of a C/EBP family.' cited in the application --- -/--	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

20 September 1995

Date of mailing of the international search report

10. 10. 95

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+ 31-70) 340-3016

Authorized officer

Cartagena y Abella,P

## INTERNATIONAL SEARCH REPORT

Inter nal Application No

PCT/US 95/06524

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. IMMUNOL., vol. 142, no. 9, 1 May 1989 pages 3134-3139, S. S. TABIBZADEH ET AL. 'Cytokine-induced production of IFN- $\gamma$ /IL-6 by freshly explanted human endometrial stromal cells.' cited in the application	
P,A	----- MOLECULAR ENDOCRINOLOGY, vol. 9, no. 4, April 1995 pages 401-412, E. CALDENHOVEN ET AL. 'Negative cross-talk between RelA and the Glucocorticoid receptor: a possible mechanism for the antiinflammatory action of glucocorticoids.' -----	

**information on patent family members**

**PCT/US 95/06524**

Form PCT/ISA/210 (patent family annex) (July 1992)



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C07K 11/00, C07H 21/04, C12P 21/02, C12N 5/10, A61K 38/43</b>		<b>A1</b>	(11) International Publication Number: <b>WO 96/30402</b>
			(43) International Publication Date: <b>3 October 1996 (03.10.96)</b>
(21) International Application Number: <b>PCT/US96/04101</b>		(74) Agents: MISROCK, S., Leslie et al.; Pennie & Edmonds, 1155 Avenue of the Americas, New York, NY 10036 (US).	
(22) International Filing Date: <b>26 March 1996 (26.03.96)</b>		(81) Designated States: AL, AM, AU, AZ, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, KZ, LK, LR, LS, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, TR, TT, UA, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(30) Priority Data: <b>411,111</b> <b>27 March 1995 (27.03.95)</b> <b>US</b>			
(60) Parent Application or Grant (63) Related by Continuation <b>US</b> <b>411,111 (CIP)</b> <b>Filed on</b> <b>27 March 1995 (27.03.95)</b>			
(71) Applicant (for all designated States except US): <b>YALE UNIVERSITY [US/US]; 451 College Street, New Haven, CT 06520 (US).</b>		<b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(72) Inventors; and (75) Inventors/Applicants (for US only): <b>XU, Tian [CN/US]; 149 Flagmarsh Road, Guilford, CT 06437 (US). TAO, Wufan [CN/US]; 32 Kenwood Lane, Branford, CT 06405 (US). WANG, Weiyi [CN/US]; 78 Edwards Street, New Haven, CT 06511 (US). ZHANG, Sheng [CN/US]; 180 Edwards Street, New Haven, CT 06511 (US). YU, Wan [CN/US]; 149 Flagmarsh Road, Guilford, CT 06437 (US).</b>			
(54) Title: <b>NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS GENES AND METHODS BASED THEREON</b>			
(57) Abstract <p>The present invention relates to a tumor suppressor gene, termed large tumor suppressor (lats), and methods for identifying tumor suppressor genes. The method provides nucleotide sequences of <i>lats</i> genes, and amino acid sequences of their encoded proteins, as well as derivatives (e.g., fragments) and analogs thereof. In a specific embodiment, the lats protein is a human protein. The invention further relates to fragments (and derivatives and analogs thereof) of lats which comprise one or more domains of a lats protein. Antibodies to lats, its derivatives and analogs, are additionally provided. Methods of production of the lats proteins, derivatives and analogs, e.g., by recombinant means, are also provided. Therapeutic and diagnostic methods and pharmaceutical compositions are provided. The invention also relates to recombinant plants and animals and methods of increasing the growth of edible plants and animals. In specific examples, isolated lats genes, from <i>Drosophila</i>, mouse, and human, and the sequences thereof, are provided.</p>			



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

**NUCLEOTIDE AND PROTEIN SEQUENCES  
OF LATS GENES AND METHODS BASED THEREON**

**1. INTRODUCTION**

This application is a continuation-in-part of  
5 copending application Serial No. 08/411,111 filed March 27,  
1995, which is incorporated by reference herein in its  
entirety.

The present invention relates to tumor suppressor  
genes, in particular to "lats" genes (large tumor suppressor)  
10 and their encoded protein products, as well as derivatives  
and analogs thereof. Production of lats proteins,  
derivatives, and antibodies is also provided. The invention  
further relates to therapeutic compositions and methods of  
diagnosis and therapy.

15

**2. BACKGROUND OF THE INVENTION**

Tumorigenesis in humans is a complex process  
involving activation of oncogenes and inactivation of tumor  
suppressor genes (Bishop, 1991, Cell 64:235-248). Tumor  
20 suppressor genes in humans have been identified through  
studies of genetic changes occurring in cancer cells (Ponder,  
1990, Trends Genet. 6:213-218; Weinberg, 1991, Science  
254:1138-1146). In *Drosophila*, tumor suppressor genes have  
been previously identified by recessive overproliferation  
25 mutations that cause late larval and pupal lethality (Gateff,  
1978, Science 200:1448-1459; Gateff and Mechler, 1989, CRC  
Crit. Rev. Oncogen 1:221-245; Bryant, 1993, Trends Cell Biol.  
3:31-35; Török et al., 1993, Genetics 135:71-80). Mutations  
of interest were identified when dissection of dead larvae  
30 and pupae revealed certain overproliferated tissues. Several  
genes identified in homozygous mutants have been cloned  
including *l(1)discs large-1(dlg)*; Woods and Bryant, 1991, Cell  
66:451-464; Woods and Bryant, 1993, Mechanisms of Development  
44:85-89), *fat* (Mahoney et al., 1991, Cell 67:853-868),  
35 *l(2)giant larvae (lgl)*. Lützelschwab et al., 1987, EMBO J.  
6:1791-1797; Jacob et al., 1987, Cell 50:215-225), *expanded*  
(*ex*; Boedigheimer and Laughon, 1993, Development

118:1291-1301; Boedigheim et al., 1993, Mechanisms of Development 44:83-84), hyperplastic discs (hyd; Mansfield et al., 1994, Developmental Biology 165:507-526) and the gene encoding the S6 ribosomal protein (Watson et al., 1992, Proc. Natl. Acad. Sci. USA 89:11302-11306; Stewart and Denell, 1993, Mol. Cell. Biol. 13:2524-2535).

Although examining homozygous mutant animals has allowed the successful identification of overproliferation mutations that cause late larval and pupal lethality, mutations that cause lethality at early developmental stages are unlikely to be recovered by this approach. The present invention solves this problem by providing a method for identifying tumor suppressor genes that does not exclude genes that when mutated cause lethality in early developmental stages, and provides genes thus identified with a fundamental role in regulation of cell proliferation.

The cessation of proliferative capacity by cells in culture is termed cellular senescence. Cellular senescence is used as an experimental model for cellular aging. Normal vertebrate cells in culture have a finite lifespan in that they undergo a characteristic maximum number of population doublings. The maximum number of population doublings that a cell can undergo inversely correlates with the age of the human donor. Cells from many human tumors are immortal cell lines when grown in tissue culture, i.e., they exhibit infinite or continuous cell growth, suggesting that overcoming senescence is part of carcinogenesis. (For the foregoing see Hubbard and Ozer, 1995, "Senescence and immortalization of human cells," in Cell Growth and Apoptosis, A Practical Approach, Ch. 12, Studzinski, G.P. (ed.), Oxford University Press Inc., New York, NY, pp. 229-248; Hubbard-Smith et al., 1992, Mol. Cell. Biol. 12:2273-2281). A comparative study of preimmortalized and immortalized human fibroblasts transformed with a defective SV40 genome has led to the suggestion that a chromosomal region at and/or distal to 6q21 plays a role in

immortalization of cells (Hubbard-Smith et al., 1992, Mol. Cell. Biol. 12:2273-2281).

Citation of references hereinabove shall not be construed as an admission that such references are prior art  
5 to the present invention.

### 3. SUMMARY OF THE INVENTION

The present invention relates to nucleotide sequences of lats genes (*Drosophila*, human, and mouse lats  
10 and lats homologs of other species), and amino acid sequences of their encoded proteins, as well as derivatives (e.g., fragments) and analogs thereof. Nucleic acids hybridizable to or complementary to the foregoing nucleotide sequences are also provided. In a specific embodiment, the lats protein is  
15 a human protein.

The invention also relates to a method of identifying tumor suppressor genes that does not exclude from identification genes that cause lethality at early developmental stages, thus overcoming the limitations of  
20 prior art methods. The method thus allows the identification of genes that regulate cell proliferation and that act at early developmental stages. The genes which thus can be identified play a fundamental role in regulation of cell proliferation such that their dysfunction (e.g., by lack of  
25 expression or mutation) leads to overproliferation and cancer.

*Lats* is a gene provided by the present invention, identified by the method of the invention, that acts to inhibit cell proliferation, and that plays a crucial role  
30 throughout development.

The invention also relates to lats derivatives and analogs of the invention which are functionally active, i.e., they are capable of displaying one or more known functional activities associated with a full-length (wild-type) lats  
35 protein. Such functional activities include but are not limited to kinase activity, antigenicity [ability to bind (or compete with lats for binding) to an anti-lats antibody],

immunogenicity (ability to generate antibody which binds to lats), and ability to bind (or compete with lats for binding) to a receptor/ligand for lats (e.g., a SH3 domain-containing protein).

5           The invention further relates to fragments (and derivatives and analogs thereof) of lats which comprise one or more domains of a lats protein.

Antibodies to lats, and lats derivatives and analogs, are additionally provided.

10           Methods of production of the lats proteins, derivatives and analogs, e.g., by recombinant means, are also provided.

The present invention also relates to therapeutic and diagnostic methods and compositions based on lats  
15 proteins and nucleic acids. Therapeutic compounds of the invention include but are not limited to lats proteins and analogs and derivatives (including fragments) thereof; antibodies thereto; nucleic acids encoding the lats proteins, analogs, or derivatives; and lats antisense nucleic acids.

20           The invention provides for treatment of disorders of overproliferation (e.g., cancer and hyperproliferative disorders) by administering compounds that promote lats activity (e.g., lats, an agonist of lats; nucleic acids that encode lats).

25           The invention also provides methods of treatment of disorders involving deficient cell proliferation (growth) or in which cell proliferation is otherwise desired (e.g., degenerative disorders, growth deficiencies, lesions, physical trauma) by administering compounds that antagonize,  
30 (inhibit) lats function (e.g., antibodies, antisense nucleic acids).

In a specific embodiment, lats function is antagonized in order to inhibit cellular senescence, *in vivo* or *in vitro*.

35           Antagonizing lats function can also be done to grow larger animals and plants, e.g., those used as food or material sources.

Animal models, diagnostic methods and screening methods for predisposition to disorders, and methods to identify lats agonists and antagonists, are also provided by the invention.

5

### 3.1. DEFINITIONS

As used herein, underscoring or italicizing the name of a gene shall indicate the gene, in contrast to its encoded protein product which is indicated by the name of the gene in the absence of any underscoring or italicizing. For example, "lats" shall mean the lats gene, whereas "lats" shall indicate the protein product of the lats gene.

### 4. DESCRIPTION OF THE FIGURES

Figure 1. Identifying overproliferation mutations in mosaic flies. (A) Although animals that are homozygous for a lethal mutation could die at an early developmental stage, mosaic flies carrying clones of cells that are homozygous for the same mutation could live. One can identify potential tumor suppressors by generating and examining clones of overproliferated mutant cells in mosaic animals. The genetic constitution of these mosaic flies is similar to the mosaicism of the tumor patients. (B) Genetic scheme. The P-element insertions carrying the FLP recombinase (hsFLP; Golic and Lindquist, 1989, Cell 59:499-509), its target site, FRT (solid arrows, Xu and Rubin, 1993, Development 117:1223-1237), the yellow<sup>+</sup> and mini-white<sup>+</sup> marker genes (y<sup>+</sup> and mini-w<sup>+</sup>, open arrows) are indicated. Mutagenized males were crossed to females to produce heterozygous embryos. Clones of cells homozygous for the induced mutations were generated in developing first-instar larvae by mitotic recombination at the FRT sites induced with the FLP recombinase. Mosaic adults were examined for overproliferated mutant patches (w<sup>-</sup>, y<sup>-</sup>). Individuals carrying clones of interest were then mated to recover the mutations of interest in the next generation (Xu and Rubin,

1993, Development 117:1223-1237; Xu and Harrison, 1994; Methods in Cell Biology 44:655-682). Clones of ommatidia derived from fast proliferating mutant cells were identified since they were larger than their darkly pigmented wt (wild-type) twin-spot clones (*mini-w<sup>+</sup>/mini-w<sup>+</sup>*).

Figure 2. Mutant phenotypes. (A) A clone of unpatterned, overproliferated *lats* mutant cells in the eye. (B) Induced at the same stage, the 93B mutant cells formed a less overproliferated clone. (C) A third instar *lats<sup>26-1</sup>* larva (right) was much larger than a wt sibling (left; at 18°C). (D) Wing discs from the larva in (C) (wt, top; *lats<sup>26-1</sup>*, bottom). (E) Dissected central nervous systems (wt, top; *lats<sup>26-1</sup>*, bottom). (F) A SEM (scanning electron microscope) view of a *lats* clone near the eye. (G) A closer view of a region in (F) showing the irregularity of the sizes and shapes of the mutant cells. (H) A plastic section of a mutant clone similar to the one in (F). Cells seem to be "budding" out of the surface to form new proliferating lobes (arrows). (I) A *lats* clone on the back. The boxed area is shown in (J). The bristles in the mutant clone are short, bent and often split (arrows). (K) A closer view of the hairs in a *lats* clone on the body showing enlarged bases and bent tips. (L) A section of a *lats* clone on the back showing extra cuticle deposits (arrows). All the mutant clones were induced with *lats<sup>1</sup>* unless stated differently.

Figure 3. Organization of the *Drosophila lats* gene. The genomic restriction map of the *lats* region is aligned with the *lats* 5.7 kb transcript unit. The direction of transcription is indicated with large arrows. The sizes of the *lats* introns are as follows: intron 1 (5.0 kb), intron 2 (5.8 kb), intron 3 (68 bp), intron 4 (63 bp), intron 5 (64 bp), intron 6 (61 bp), intron 7 (62 bp). The genomic DNA from +7.5 (*Bgl*II) to -4.2 (*Eco*RI) was used to screen a total imaginal disc cDNA library, which isolated three groups of cDNAs: *lats*, T1, T2. The introns in the T2 transcript are not labeled. Only parts of the *zfh-1* (Fortini et al., 1991, Mechan. Dev. 34:113-122) and T1 transcripts are

indicated. The locations of the P-element insertion (*lats<sup>PI</sup>*), the deletions in the five excision alleles (*lats<sup>e7-2, e78, e100, e119, e148</sup>*) and in *lats<sup>a1</sup>*, *lats<sup>a4</sup>* are indicated at the bottom. The slash indicates a gap in the genomic map. Restriction sites:

5 *EcoRI* (small open arrow), *BglII* (open box) and *BamHI* (open circle). The *BglII* site at the -0.5 position of the CLT-52 clone is not present in other genomic DNA. A scale is labeled under the restriction map.

Figure 4. RNA blot analysis of the *Drosophila lats* mRNA. Five µg of poly(A)<sup>+</sup> RNA isolated from various developmental stages was separated on a 1% agarose gel, and hybridized with <sup>32</sup>P-labeled 5' end 1 kb probe from the *Drosophila lats* cDNA. E0-2 hrs, E2-4 hrs, E4-6 hrs, E6-8 hrs, E8-16 hrs and E16-24 hrs indicate the age of the embryos in hours. RNA from first, second and third instar larvae is denoted by L1, L2, and L3, respectively. The numbers and arrows on the right correspond to the size and location of the RNA standards. A 5.7 kb RNA was found in all the developmental stages, whereas a 4.7 kb RNA was predominantly present in 0 to 4 hour old embryos. The blot was also hybridized with DNA from the ribosomal protein gene, RNA1.

Figure 5. Composite cDNA sequence of the *Drosophila lats* gene. The entire cDNA sequence (SEQ ID NO:1) corresponding to the 5.7 kb *lats* RNA is shown. This nucleotide sequence is a composite of two cDNA clones (nucleotide 1-191 from cDNA 9 and the rest from cDNA A2). The sequence of the corresponding genomic DNA has been determined and is identical to the cDNA sequence except where indicated (above the cDNA sequence). The predicted amino acid sequence (SEQ ID NO:2) is shown below the cDNA sequence. The opa repeat is indicated by the heavy bar. The location of the putative SH3 binding site and the RERDQ peptides are designated by dashed lines. The two sites that match the polyadenylation signal consensus sequence are underlined. The second site is located at 12 bp away from the 3' end of the cDNA. The locations of the introns are indicated by vertical arrows. The underlined 141 bp sequence at the 3'



end of the *lats* transcript is identical to the 5' end untranslated sequence of the class I transcript of the *Drosophila* phospholipase C gene, *plc-21*. The location of the 446 bp deletion in the *lats<sup>al</sup>* allele is also indicated.

5                    Figure 6. Schematic of the *Drosophila lats* predicted protein (SEQ ID NO:2) and the related proteins (A) and sequence comparison of the proteins homologous to *lats* (B). In Fig. 6A, solid, hatched, open and shaded boxes denote putative SH3 binding site, opa repeat, RERDQ peptide  
10 and kinase domain in the *lats* protein, respectively. The Dbf20, Dbf2 and COT-1 proteins are illustrated at the bottom. The regions that are homologous to *lats* are indicated by shaded boxes. The degrees of sequence similarity (percentage of identical sequences inside parentheses; percentage of  
15 identical or conservative substitutions outside parentheses) between *lats* and the three related proteins are indicated above the corresponding regions of these proteins. In Fig. 6B, the carboxy-terminal half of *lats* is compared to the six most related proteins that are revealed by blastp (a software  
20 program that searches for protein sequence homologies) search as of Sept. 1, 1994. *Neurospora cot-1* (SEQ ID NO:11); tobacco PKTL7 (SEQ ID NO:12); common ice plant protein kinase (SEQ ID NO:13); spinach protein kinase (SEQ ID NO:14); yeast Dbf-20 (SEQ ID NO:15); yeast Dbf2 (SEQ ID NO:16). Amino acid  
25 residues identical to *lats* are highlighted. Numbers at the beginning of every sequence refer to the position of that amino acid within the total protein sequence. The boundary of the kinase domain is defined according to Hanks et al. (1988, Science 241:42-52). The location of a region of about  
30 40 amino acid residues that is not conserved among the proteins is indicated by the heavy bar above the sequence. The sequence of PKTL7 from tobacco, *Nicotiana tabacum*, was submitted to Genbank by Huang, Y. (X71057). Both the sequence of the protein kinase from spinach, *Spinacia oleracea*, and  
35 the sequence of the protein kinase from common ice plant, *Mesembryanthemum crystallinum*, were submitted to Genbank by

Baur, B., Winter, K., Fischer, K. and Dietz, K. (Z30329 and Z30330).

Figur 7. cDNA sequence (SEQ ID NO:5) and deduced prot in sequence (SEQ ID NO:6) of a mouse lats homolog, 5 m-lats.

Figure 8. cDNA sequence (SEQ ID NO:7) and deduced protein sequence (SEQ ID NO:8) of a mouse lats homolog, m-lats2.

Figure 9. cDNA sequence (SEQ ID NO:3) and deduced 10 protein sequence (SEQ ID NO:4) of a human lats homolog, h-lats.

Figure 10. Schematic diagram of plasmid PBS(KS)-h-lats, containing the full length coding sequence of the h-lats cDNA.

15 Figure 11. Alignment of the h-lats protein sequence (SEQ ID NO:4) (upper case letters) with the m-lats protein sequence (SEQ ID NO:6) (lower case letters). A dot indicates amino acid identity; a dash indicates a deletion relative to the sequence on the line above. The amino- 20 terminal portion of the m-lats protein is not shown due to the missing 5' end of the m-lats cDNA coding region.

Figure 12. Alignment of the h-lats protein sequence (SEQ ID NO:4) (upper case letters) with the m-lats2 protein sequence (SEQ ID NO:8) (lower case letters). A dot 25 indicates amino acid identity; a dash indicates a deletion relative to the sequence on the line above. The amino-terminal portion of the m-lats2 protein is not shown due to the missing 5' end of the m-lats2 cDNA coding region.

Figure 13. Alignment of the h-lats protein 30 sequence (SEQ ID NO:4) (upper case letters) with the *Drosophila* lats protein sequence (SEQ ID NO:2) (lower case letters). A dot indicates amino acid identity; a dash indicates a deletion relative to the sequence on the line above. Insertions in the *Drosophila* sequence relative to the 35 human sequence are indicated below the sequence line. Conserved domains are indicated. LSD2 = lats split domain 2; LSD2a = LSD2 anterior portion; LSD2p = LSD2 posterior

portion. The putative SH3-binding domain and the kinase domain are shown. LSD1 = lats split domain 1; LSD1a = LSD1 anterior portion; LSD1p = LSD1 posterior portion. LFD = lats flanking domain. LCD1 = lats C-terminal domain 1; LCD2 = lats C-terminal domain 2; LCD3 = lats C-terminal domain 3.

Figure 14. Schematic diagram of plasmid pCaSpeR-hs-h-lats, an expression vector containing the full length coding sequence of the h-lats cDNA.

Figure 15. Northern blot analysis of h-lats expression in normal human tissues. A  $^{32}\text{P}$ -labeled BamHI fragment of h-lats was used as a probe for hybridization to polyA<sup>+</sup> RNA from the normal human fetal and adult tissues indicated for each lane. The positions of standard molecular weight markers are shown at right. The positions of the h-lats RNA and of  $\beta$ -actin RNA (used as a standard) are shown.

## 5. DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to nucleotide sequences of lats genes, and amino acid sequences of their encoded proteins. The invention further relates to fragments and other derivatives, and analogs, of lats proteins. Nucleic acids encoding such fragments or derivatives are also within the scope of the invention. The invention provides lats genes and their encoded proteins of many different species. The lats genes of the invention include *Drosophila*, human, and mouse lats and related genes (homologs) in other species. In specific embodiments, the lats genes and proteins are from vertebrates, or more particularly, mammals. In a preferred embodiment of the invention, the lats genes and proteins are of human origin. Production of the foregoing proteins and derivatives, e.g., by recombinant methods, is provided.

The invention also relates to a method of identifying tumor suppressor genes that does not exclude from identification genes that cause lethality at early developmental stages, thus overcoming the limitations of prior art methods. The method thus allows the identification

f genes that regulate cell proliferation and that act at early developmental stages. The genes which thus can be identified play a fundamental role in regulation of cell proliferation such that their dysfunction (e.g., due to lack of expression or mutation) leads to overproliferation and cancer.

Lats is a gene provided by the present invention, identified by the method of the invention, that acts to inhibit cell proliferation, and that plays a crucial role throughout development.

The invention also relates to lats derivatives and analogs of the invention which are functionally active, i.e., they are capable of displaying one or more known functional activities associated with a full-length (wild-type) lats protein. Such functional activities include but are not limited to kinase activity, antigenicity [ability to bind (or compete with lats for binding) to an anti-lats antibody], immunogenicity (ability to generate antibody which binds to lats), ability to bind (or compete with lats for binding) to an SH3-domain-containing protein or other ligand, ability to inhibit cell proliferation, tumor inhibition, etc.

The invention further relates to fragments (and derivatives and analogs thereof) of lats which comprise one or more domains of the lats protein.

Antibodies to lats, its derivatives and analogs, are additionally provided.

The present invention also relates to therapeutic and diagnostic methods and compositions based on lats proteins and nucleic acids and anti-lats antibodies. The invention provides for treatment of disorders of overproliferation (e.g., cancer and hyperproliferative disorders) by administering compounds that promote lats activity (e.g., lats proteins and functionally active analogs and derivatives (including fragments) thereof; nucleic acids encoding the lats proteins, analogs, or derivatives, agonists of lats).

The invention also provides methods of treatment of disorders involving deficient cell proliferation or in which cell proliferation (growth) is otherwise desirable (e.g., growth deficiencies, degenerative disorders, lesions, physical trauma) by administering compounds that antagonize, or inhibit, *lats* function (e.g., antibodies, *lats* antisense nucleic acids, *lats* derivatives that are dominant-negative protein kinases).

In a specific embodiment, *lats* function is antagonized in order to inhibit cellular senescence, *in vivo* or *in vitro*.

Inhibition of *lats* function can also be done to grow larger farm animals and plants.

Animal models, diagnostic methods and screening methods for predisposition to disorders are also provided by the invention.

The invention is illustrated by way of examples *infra* which disclose, *inter alia*, the cloning and characterization of *D. melanogaster lats* (Section 6); the cloning and characterization of mouse and human *lats* homologs (Section 7); the sequence and domain conservation among the *lats* homologs (Section 8); the functional interchangeability of the human and *Drosophila lats* homologs (Section 9); and the differentially decreased expression of human *lats* in human tumor cell lines (Section 10).

For clarity of disclosure, and not by way of limitation, the detailed description of the invention is divided into the subsections which follow.

### 5.1. ISOLATION OF THE *LATS* GENES

The invention relates to the nucleotide sequences of *lats* nucleic acids. In specific embodiments, *lats* nucleic acids comprise the cDNA sequences of SEQ ID NO:1, 3, 5, or 7, or the coding regions thereof, or nucleotide sequences encoding a *lats* protein (e.g., a protein having the sequence of SEQ ID NO:2, 4, 6, or 8). The invention provides purified nucleic acids consisting of at least 8 nucleotides (i.e., a

hybridizable portion) of a *lats* sequence; in other embodiments, the nucleic acids consist of at least 25 (continuous) nucleotides, 50 nucleotides, 100 nucleotides, 150 nucleotides, or 200 nucleotides of a *lats* sequence, or a full-length *lats* coding sequence. In another embodiment, the nucleic acids are smaller than 35, 200 or 500 nucleotides in length. Nucleic acids can be single or double stranded. The invention also relates to nucleic acids hybridizable to or complementary to the foregoing sequences. In specific aspects, nucleic acids are provided which comprise a sequence complementary to at least 10, 25, 50, 100, or 200 nucleotides or the entire coding region of a *lats* gene. In a specific embodiment, a nucleic acid which is hybridizable to a *lats* nucleic acid (e.g., having sequence SEQ ID NO:3 or 7), or to a nucleic acid encoding a *lats* derivative, under conditions of low stringency is provided. By way of example and not limitation, procedures using such conditions of low stringency are as follows (see also Shilo and Weinberg, 1981, Proc. Natl. Acad. Sci. USA 78:6789-6792): Filters containing DNA are pretreated for 6 h at 40°C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 µg/ml denatured salmon sperm DNA. Hybridizations are carried out in the same solution with the following modifications: 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 µg/ml salmon sperm DNA, 10% (wt/vol) dextran sulfate, and 5-20 X 10<sup>6</sup> cpm <sup>32</sup>P-labeled probe is used. Filters are incubated in hybridization mixture for 18-20 h at 40°C, and then washed for 1.5 h at 55°C in a solution containing 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS. The wash solution is replaced with fresh solution and incubated an additional 1.5 h at 60°C. Filters are blotted dry and exposed for autoradiography. If necessary, filters are washed for a third time at 65-68°C and reexposed to film. Other conditions of low stringency which may be used are well known in the art (e.g., as employed for cross-species hybridizations).

In another specific embodiment, a nucleic acid which is hybridizable to a lats nucleic acid under conditions of high stringency is provided. By way of example and not limitation, procedures using such conditions of high stringency are as follows: Prehybridization of filters containing DNA is carried out for 8 h to overnight at 65°C in buffer composed of 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 µg/ml denatured salmon sperm DNA. Filters are hybridized for 48 h at 65°C in prehybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 X 10<sup>6</sup> cpm of <sup>32</sup>P-labeled probe. Washing of filters is done at 37°C for 1 h in a solution containing 2X SSC, 0.01% PVP, 0.01% Ficoll, and 0.01% BSA. This is followed by a wash in 0.1X SSC at 50°C for 45 min before autoradiography. Other conditions of high stringency which may be used are well known in the art.

In another specific embodiment, a nucleic acid, which is hybridizable to a lats nucleic acid under conditions of moderate stringency is provided (see, e.g., Section 7.2).

20 Nucleic acids encoding derivatives and analogs of lats proteins (see Sections 5.6 and 5.6.1), and lats antisense nucleic acids (see Section 5.8.2.2.1) are additionally provided. As is readily apparent, as used herein, a "nucleic acid encoding a fragment or portion of a lats protein" shall be construed as referring to a nucleic acid encoding only the recited fragment or portion of the lats protein and not the other contiguous portions of the lats protein as a continuous sequence.

Fragments of lats nucleic acids comprising regions conserved between (with homology to) other lats nucleic acids, of the same or different species, are also provided. Nucleic acids encoding one or more lats domains are provided.

Specific embodiments for the cloning of a lats gene, presented as a particular example but not by way of limitation, follows:

For expression cloning (a technique commonly known in the art), an expression library is constructed by methods

known in the art. For example, mRNA (e.g., human) is isolated, cDNA is made and ligated into an expression vector (e.g., a bacteriophag derivative) such that it is capable of being expressed by the host cell into which it is then  
5 introduced. Various screening assays can then be used to select for the expressed lats product. In one embodiment, anti-lats antibodies can be used for selection.

In another embodiment, polymerase chain reaction (PCR) is used to amplify the desired sequence in a genomic or  
10 cDNA library, prior to selection. Oligonucleotide primers representing known lats sequences can be used as primers in PCR. In a preferred aspect, the oligonucleotide primers represent at least part of the lats conserved segments of strong homology between lats of different species (e.g.,  
15 LCD1, LCD2, kinase domain, LFD, SH3 binding domain, LSD1, and LSD2 domains; see, e.g., Section 8 *infra*.) The synthetic oligonucleotides may be utilized as primers to amplify by PCR sequences from a source (RNA or DNA), preferably a cDNA library, of potential interest. PCR can be carried out,  
20 e.g., by use of a Perkin-Elmer Cetus thermal cyclor and Taq polymerase (Gene Amp™). The DNA being amplified can include mRNA or cDNA or genomic DNA from any eukaryotic species. One can choose to synthesize several different degenerate primers, for use in the PCR reactions. It is also possible  
25 to vary the stringency of hybridization conditions used in priming the PCR reactions, to allow for greater or lesser degrees of nucleotide sequence similarity between the known lats nucleotide sequence and the nucleic acid homolog being isolated. For cross species hybridization, low stringency  
30 conditions are preferred. For same species hybridization, moderately stringent conditions are preferred. After successful amplification of a segment of a lats homolog, that segment may be molecularly cloned and sequenced, and utilized as a probe to isolate a complete cDNA or genomic clone.  
35 This, in turn, will permit the determination of the gene's complete nucleotide sequence, the analysis of its expression, and the production of its protein product for functional



analysis, as described *infra*. In this fashion, additional genes encoding latex proteins and latex analogs may be identified.

The above-methods are not meant to limit the following general description of methods by which clones of latex may be obtained.

Any eukaryotic cell potentially can serve as the nucleic acid source for the molecular cloning of the latex gene. The nucleic acid sequences encoding latex can be isolated from vertebrate, mammalian, human, porcine, bovine, feline, avian, equine, canine, as well as additional primate sources, insects, plants, etc. The DNA may be obtained by standard procedures known in the art from cloned DNA (e.g., a DNA "library"), by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell. (See, for example, Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II.) Clones derived from genomic DNA may contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will contain only exon sequences. Whatever the source, the gene should be molecularly cloned into a suitable vector for propagation of the gene.

In the molecular cloning of the gene from genomic DNA, DNA fragments are generated, some of which will encode the desired gene. The DNA may be cleaved at specific sites using various restriction enzymes. Alternatively, one may use DNase in the presence of manganese to fragment the DNA, or the DNA can be physically sheared, as for example, by sonication. The linear DNA fragments can then be separated according to size by standard techniques, including but not limited to, agarose and polyacrylamide gel electrophoresis and column chromatography.

Once the DNA fragments are generated, identification of the specific DNA fragment containing the

d sir d gene may b accomplished in a number of ways. For example, if an amount of a portion of a lats (of any species) gen or its specific RNA, or a fragment thereof (see Section 5.6), is available and can be purified and labeled, the 5 generated DNA fragments may be screened by nucleic acid hybridization to the labeled probe (Benton, W. and Davis, R., 1977, Science 196:180; Grunstein, M. And Hogness, D., 1975, Proc. Natl. Acad. Sci. U.S.A. 72:3961). Such a procedure is presented by way of example in Section 7 *infra*. Those DNA 10 fragments with substantial homology to the probe will hybridize. It is also possible to identify the appropriate fragment by restriction enzyme digestion(s) and comparison of fragment sizes with those expected according to a known restriction map if such is available. Further selection can 15 be carried out on the basis of the properties of the gene. Alternatively, the presence of the gene may be detected by assays based on the physical, chemical, or immunological properties of its expressed product. For example, cDNA clones, or DNA clones which hybrid-select the proper mRNAs, 20 can be selected which produce a protein that, e.g., has similar or identical electrophoretic migration, isoelectric focusing behavior, proteolytic digestion maps, kinase activity, inhibition of cell proliferation activity, substrate binding activity, or antigenic properties as known 25 for lats. If an antibody to lats is available, the lats protein may be identified by binding of labeled antibody to the putatively lats synthesizing clones, in an ELISA (enzyme-linked immunosorbent assay)-type procedure.

The lats gene can also be identified by mRNA 30 selection by nucleic acid hybridization followed by *in vitro* translation. In this procedure, fragments are used to isolate complementary mRNAs by hybridization. Such DNA fragments may represent available, purified lats DNA of another species (e.g., *Drosophila*, mouse, human).

35 Immunoprecipitation analysis or functional assays (e.g., aggregation ability *in vitro*; binding to receptor; se *infra*) of th *in vitro* translation products of the isolated products

of the isolated mRNAs identifies the mRNA and, therefore, the complementary DNA fragments that contain the desired sequences. In addition, specific mRNAs may be selected by adsorption of polysomes isolated from cells to immobilized antibodies specifically directed against lats protein. A radiolabelled lats cDNA can be synthesized using the selected mRNA (from the adsorbed polysomes) as a template. The radiolabelled mRNA or cDNA may then be used as a probe to identify the lats DNA fragments from among other genomic DNA fragments.

Alternatives to isolating the lats genomic DNA include, but are not limited to, chemically synthesizing the gene sequence itself from a known sequence or making cDNA to the mRNA which encodes the lats protein. For example, RNA for cDNA cloning of the lats gene can be isolated from cells which express lats. Other methods are possible and within the scope of the invention.

The identified and isolated gene can then be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art may be used. Possible vectors include, but are not limited to, plasmids or modified viruses, but the vector system must be compatible with the host cell used. Such vectors include, but are not limited to, bacteriophages such as lambda derivatives, or plasmids such as PBR322 or pUC plasmid derivatives or the Bluescript vector (Stratagene). The insertion into a cloning vector can, for example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. However, if the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules may be enzymatically modified. Alternatively, any site desired may be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences. In an alternative method, the cleaved vector and lats gene may be modified by

homopolymERIC tailing. Recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, etc., so that many copies of the gene sequence are generated.

5 In an alternative method, the desired gene may be identified and isolated after insertion into a suitable cloning vector in a "shot gun" approach. Enrichment for the desired gene, for example, by size fractionization, can be done before insertion into the cloning vector.

10 In specific embodiments, transformation of host cells with recombinant DNA molecules that incorporate the isolated lats gene, cDNA, or synthesized DNA sequence enables generation of multiple copies of the gene. Thus, the gene may be obtained in large quantities by growing transformants, 15 isolating the recombinant DNA molecules from the transformants and, when necessary, retrieving the inserted gene from the isolated recombinant DNA.

The lats sequences provided by the instant invention include those nucleotide sequences encoding 20 substantially the same amino acid sequences as found in native lats proteins, and those encoded amino acid sequences with functionally equivalent amino acids, as well as those encoding other lats derivatives or analogs, as described in Sections 5.6 and 5.6.1 *infra* for lats derivatives and 25 analogs.

## 5.2. EXPRESSION OF THE LATS GENES

The nucleotide sequence coding for a lats protein or a functionally active analog or fragment or other 30 derivative thereof (see Section 5.6), can be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. The necessary transcriptional and translational signals can also be 35 supplied by the native lats gene and/or its flanking regions. A variety of host-vector systems may be utilized to express the protein-coding sequence. These include but are not

limited to mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus); microorganisms such as yeast containing yeast vectors, or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their strengths and specificities. Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be used. In specific embodiments, the human lats gene is expressed, or a sequence encoding a functionally active portion of human lats. In yet another embodiment, a fragment of lats comprising a domain of the lats protein is expressed.

Any of the methods previously described for the insertion of DNA fragments into a vector may be used to construct expression vectors containing a chimeric gene consisting of appropriate transcriptional/translational control signals and the protein coding sequences. These methods may include *in vitro* recombinant DNA and synthetic techniques and *in vivo* recombinants (genetic recombination). Expression of nucleic acid sequence encoding a lats protein or peptide fragment may be regulated by a second nucleic acid sequence so that the lats protein or peptide is expressed in a host transformed with the recombinant DNA molecule. For example, expression of a lats protein may be controlled by any promoter/enhancer element known in the art. In a specific embodiment, the promoter is not a native lats gene promoter. Promoters which may be used to control lats expression include, but are not limited to, the SV40 early promoter region (Berne and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, et al., 1980, Cell 22:787-797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the regulatory sequences of the metallothionein gene (Brinster et al., 1982, Nature 296:39-42); prokaryotic expression vectors such as the  $\beta$ -lactamase promoter (Villa-Komaroff, et al.,

1978, *Proc. Natl. Acad. Sci. U.S.A.* 75:3727-3731), or the *tac* promoter (D Bo r, et al., 1983, *Proc. Natl. Acad. Sci. U.S.A.* 80:21-25); see also "Useful proteins from recombinant bacteria" in *Scientific American*, 1980, 242:74-94; plant  
5 expression vectors comprising the nopaline synthetase promoter region (Herrera-Estrella et al., *Nature* 303:209-213) or the cauliflower mosaic virus 35S RNA promoter (Gardner, et al., 1981, *Nucl. Acids Res.* 9:2871), and the promoter of the photosynthetic enzyme ribulose biphosphate carboxylase  
10 (Herrera-Estrella et al., 1984, *Nature* 310:115-120); promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADC (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter, and the following animal transcriptional control  
15 regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift et al., 1984, *Cell* 38:639-646; Ornitz et al., 1986, *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409; MacDonald, 1987,  
20 *Hepatology* 7:425-515); insulin gene control region which is active in pancreatic beta cells (Hanahan, 1985, *Nature* 315:115-122), immunoglobulin gene control region which is active in lymphoid cells (Grosschedl et al., 1984, *Cell* 38:647-658; Adames et al., 1985, *Nature* 318:533-538;  
25 Alexander et al., 1987, *Mol. Cell. Biol.* 7:1436-1444), mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder et al., 1986, *Cell* 45:485-495), albumin gene control region which is active in liver (Pinkert et al., 1987, *Genes and Devel.*  
30 1:268-276), alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., 1985, *Mol. Cell. Biol.* 5:1639-1648; Hammer et al., 1987, *Science* 235:53-58; alpha 1-antitrypsin gene control region which is active in the liver (Kelsey et al., 1987, *Genes and Devel.* 1:161-171), beta-  
35 globin gene control region which is active in myeloid cells (Mogham et al., 1985, *Nature* 315:338-340; Kollias et al., 1986, *Cell* 46:89-94; myelin basic protein gene control region

which is active in oligodendrocyte cells in the brain (Radha et al., 1987, Cell 48:703-712); myosin light chain-2 gene control region which is active in skeletal muscle (Sani, 1985, Nature 314:283-286), and gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason et al., 1986, Science 234:1372-1378).

In a specific embodiment, a vector is used that comprises a promoter operably linked to a *lats*-encoding nucleic acid, one or more origins of replication, and, optionally, one or more selectable markers (e.g., an antibiotic resistance gene).

In a specific embodiment, an expression construct is made by subcloning a *lats* coding sequence into the *EcoRI* restriction site of each of the three pGEX vectors (Glutathione S-Transferase expression vectors; Smith and Johnson, 1988, Gene 7:31-40). This allows for the expression of the *lats* protein product from the subclone in the correct reading frame.

Expression vectors containing *lats* gene inserts can be identified by three general approaches: (a) nucleic acid hybridization, (b) presence or absence of "marker" gene functions, and (c) expression of inserted sequences. In the first approach, the presence of a *lats* gene inserted in an expression vector can be detected by nucleic acid hybridization using probes comprising sequences that are homologous to an inserted *lats* gene. In the second approach, the recombinant vector/host system can be identified and selected based upon the presence or absence of certain "marker" gene functions (e.g., thymidine kinase activity, resistance to antibiotics, transformation phenotype, occlusion body formation in baculovirus, etc.) caused by the insertion of a *lats* gene in the vector. For example, if the *lats* gene is inserted within the marker gene sequence of the vector, recombinants containing the *lats* insert can be identified by the absence of the marker gene function. In the third approach, recombinant expression vectors can be identified by assaying the *lats* product expressed by the

r combinant. Such assays can be based, for example, on the physical or functional properties of the lats protein in in vitro assay systems, e.g., kinase activity, binding with anti-lats antibody, inhibition of cell proliferation.

5           Once a particular recombinant DNA molecule is identified and isolated, several methods known in the art may be used to propagate it. Once a suitable host system and growth conditions are established, recombinant expression vectors can be propagated and prepared in quantity. As  
10 previously explained, the expression vectors which can be used include, but are not limited to, the following vectors or their derivatives: human or animal viruses such as vaccinia virus or adenovirus; insect viruses such as baculovirus; yeast vectors; bacteriophage vectors (e.g.,  
15 lambda), and plasmid and cosmid DNA vectors, to name but a few.

          In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific  
20 fashion desired. Expression from certain promoters can be elevated in the presence of certain inducers; thus, expression of the genetically engineered lats protein may be controlled. Furthermore, different host cells have characteristic and specific mechanisms for the translational  
25 and post-translational processing and modification (e.g., glycosylation, phosphorylation of proteins. Appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the foreign protein expressed. For example, expression in a bacterial system can be used to  
30 produce an unglycosylated core protein product. Expression in yeast will produce a glycosylated product. Expression in mammalian cells can be used to ensure "native" glycosylation of a heterologous protein. Furthermore, different vector/host expression systems may effect processing  
35 reactions to different extents.

          In other specific embodiments, the lats protein, fragment, analog, or derivative may be expressed as a fusion,



r chimeric protein product (comprising the protein, fragment, analog, or derivative joined via a peptide bond to a heterologous protein sequence (of a different protein)). Such a chimeric product can be made by ligating the  
5 appropriate nucleic acid sequences encoding the desired amino acid sequences to each other by methods known in the art, in the proper coding frame, and expressing the chimeric product by methods commonly known in the art. Alternatively, such a chimeric product may be made by protein synthetic techniques,  
10 e.g., by use of a peptide synthesizer.

Both cDNA and genomic sequences can be cloned and expressed.

### 5.3. IDENTIFICATION AND PURIFICATION 15 OF THE LATS GENE PRODUCTS

In particular aspects, the invention provides amino acid sequences of lats, preferably human lats, and fragments and derivatives thereof which comprise an antigenic determinant (i.e., can be recognized by an antibody) or which  
20 are otherwise functionally active, as well as nucleic acid sequences encoding the foregoing. "Functionally active" lats material as used herein refers to that material displaying one or more known functional activities associated with a full-length (wild-type) lats protein, e.g., kinase activity,  
25 inhibition of cell proliferation, tumor inhibition, binding to an SH3-domain, binding to a lats substrate or lats binding partner, antigenicity (binding to an anti-lats antibody), immunogenicity, etc.

In specific embodiments, the invention provides  
30 fragments of a lats protein consisting of at least 6 amino acids, 10 amino acids, 50 amino acids, or of at least 75 amino acids. In other embodiments, the proteins comprise or consist essentially of a lats carboxy (C)-terminal domain 3 (LCD3), lats C-terminal domain 2 (LCD2), lats C-terminal  
35 domain 1 (LCD1), kinase domain, kinase subdomains, lats flanking domain (amino-terminal to the kinase domain), lats split domain 1 (LSD1), lats split domain 2 (LSD2),

SH3-binding domain, and opa r peat domain (s e Section 8 *infra*), or any combination of the foregoing, of a lats protein. Fragments, or proteins comprising fragments, lacking some or all of the foregoing regions of a lats  
5 protein are also provided. Nucleic acids encoding the foregoing are provided.

Once a recombinant which expresses the lats gene sequence is identified, the gene product can be analyzed. This is achieved by assays based on the physical or  
10 functional properties of the product, including radioactive labelling of the product followed by analysis by gel electrophoresis, immunoassay, etc.

Once the lats protein is identified, it may be isolated and purified by standard methods including  
15 chromatography (e.g., ion exchange, affinity, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins. The functional properties may be evaluated using any suitable assay (see Section 5.7).

20 Alternatively, once a lats protein produced by a recombinant is identified, the amino acid sequence of the protein can be deduced from the nucleotide sequence of the chimeric gene contained in the recombinant. As a result, the protein can be synthesized by standard chemical methods known  
25 in the art (e.g., see Hunkapiller, M., et al., 1984, Nature 310:105-111).

In another alternate embodiment, native lats proteins can be purified from natural sources, by standard methods such as those described above (e.g., immunoaffinity  
30 purification).

In a specific embodiment of the present invention, such lats proteins, whether produced by recombinant DNA techniques or by chemical synthetic methods or by purification of native proteins, include but are not limited  
35 to those containing, as a primary amino acid sequence, all or part of the amino acid s qu nce substantially as depicted in Figure 9 (SEQ ID NO:4), as well as fragm nts and other

derivativ s, and analogs thereof, including proteins homologous thereto.

#### 5.4. STRUCTURE OF THE LATS GENE AND PROTEIN

5           The structure of the lats gene and protein can be analyzed by various methods known in the art.

##### 5.4.1. GENETIC ANALYSIS

          The cloned DNA or cDNA corresponding to the lats  
10 gene can be analyzed by methods including but not limited to Southern hybridization (Southern, E.M., 1975, J. Mol. Biol. 98:503-517), Northern hybridization (see e.g., Freeman et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:4094-4098), restriction endonuclease mapping (Maniatis, T., 1982, Molecular Cloning, A Laboratory, Cold Spring Harbor, New York), and DNA sequence analysis. Polymerase chain reaction (PCR; U.S. Patent Nos. 4,683,202, 4,683,195 and 4,889,818; Gyllenstein et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7652-7656; Ochman et al., 1988, Genetics 120:621-623; Loh  
20 et al., 1989, Science 243:217-220) followed by Southern hybridization with a lats-specific probe can allow the detection of the lats gene in DNA from various cell types. Methods of amplification other than PCR are commonly known and can also be employed. In one embodiment, Southern  
25 hybridization can be used to determine the genetic linkage of lats. Northern hybridization analysis can be used to determine the expression of the lats gene. Various cell types, at various states of development or activity can be tested for lats expression. The stringency of the  
30 hybridization conditions for both Southern and Northern hybridization can be manipulated to ensure detection of nucleic acids with the desired degree of relatedness to the specific lats probe used. Modifications of these methods and other methods commonly known in the art can be used.  
35           Restriction endonuclease mapping can be used to roughly determine the genetic structure of the lats gene.

Restriction maps derived by restriction endonuclease cleavage can be confirmed by DNA sequence analysis.

DNA sequence analysis can be performed by any techniques known in the art, including but not limited to the method of Maxam and Gilbert (1980, Meth. Enzymol. 65:499-560), the Sanger dideoxy method (Sanger, F., et al., 1977, Proc. Natl. Acad. Sci. U.S.A. 74:5463), the use of T7 DNA polymerase (Tabor and Richardson, U.S. Patent No. 4,795,699), or use of an automated DNA sequencer (e.g., Applied Biosystems, Foster City, CA).

#### 5.4.2. PROTEIN ANALYSIS

The amino acid sequence of the lats protein can be derived by deduction from the DNA sequence, or alternatively, by direct sequencing of the protein, e.g., with an automated amino acid sequencer.

The lats protein sequence can be further characterized by a hydrophilicity analysis (Hopp, T. and Woods, K., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:3824). A hydrophilicity profile can be used to identify the hydrophobic and hydrophilic regions of the lats protein and the corresponding regions of the gene sequence which encode such regions.

Secondary, structural analysis (Chou, P. and Fasman, G., 1974, Biochemistry 13:222) can also be done, to identify regions of lats that assume specific secondary structures.

Manipulation, translation, and secondary structure prediction, open reading frame prediction and plotting, as well as determination of sequence homologies, can also be accomplished using computer software programs available in the art.

Other methods of structural analysis can also be employed. These include but are not limited to X-ray crystallography (Engstrom, A., 1974, Biochem. Exp. Biol. 11:7-13) and computer modeling (Fletterick, R. and Zoller, M. (ds.), 1986, Computer Graphics and Molecular Modeling, in

Current Communications in Molecular Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

5.5. GENERATION OF ANTIBODIES TO LATS  
PROTEINS AND DERIVATIVES THEREOF

5 According to the invention, lats protein, its fragments or other derivatives, or analogs thereof, may be used as an immunogen to generate antibodies which immunospecifically bind such an immunogen. Such antibodies include but are not limited to polyclonal, monoclonal, 10 chimeric, single chain, Fab fragments, and an Fab expression library. In a specific embodiment, antibodies to a human lats protein are produced. In another embodiment, antibodies to a domain (e.g., the SH3-binding domain) of a lats protein are produced. In a specific embodiment, fragments of a lats 15 protein identified as hydrophilic are used as immunogens for antibody production.

Various procedures known in the art may be used for the production of polyclonal antibodies to a lats protein or derivative or analog. In a particular embodiment, rabbit 20 polyclonal antibodies to an epitope of a lats protein encoded by a sequence of SEQ ID NOS:2, 4, 6 or 8, or a subsequence thereof, can be obtained. For the production of antibody, various host animals can be immunized by injection with the native lats protein, or a synthetic version, or derivative 25 (e.g., fragment) thereof, including but not limited to rabbits, mice, rats, etc. Various adjuvants may be used to increase the immunological response, depending on the host species, and including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, 30 surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and 35 corynebacterium parvum.

For preparation of monoclonal antibodies directed toward a lats protein sequence or analog thereof, any

technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the hybridoma technique originally developed by Kohler and Milstein (1975, Nature 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals utilizing recent technology (PCT/US90/02545). According to the invention, human antibodies may be used and can be obtained by using human hybridomas (Cote et al., 1983, Proc. Natl. Acad. Sci. U.S.A. 80:2026-2030) or by transforming human B cells with EBV virus in vitro (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, pp. 77-96). In fact, according to the invention, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci. U.S.A. 81:6851-6855; Neuberger et al., 1984, Nature 312:604-608; Takeda et al., 1985, Nature 314:452-454) by splicing the genes from a mouse antibody molecule specific for rats together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention.

According to the invention, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce rats-specific single chain antibodies. An additional embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries (Huse et al., 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for rats proteins, derivatives, or analogs.

Antibody fragments which contain the idiotype of the molecule can be generated by known techniques. For example, such fragments include but are not limited to: the

F(ab')<sub>2</sub> fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragment, the Fab fragments which can be generated by  
5 treating the antibody molecule with papain and a reducing agent, and Fv fragments.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, e.g. ELISA (enzyme-linked immunosorbent assay). For  
10 example, to select antibodies which recognize a specific domain of a lats protein, one may assay generated hybridomas for a product which binds to a lats fragment containing such domain. For selection of an antibody that specifically binds a first lats homolog but which does not specifically bind a  
15 different lats homolog, one can select on the basis of positive binding to the first lats homolog and a lack of binding to the second lats homolog.

Antibodies specific to a domain of a lats protein are also provided.

20 The foregoing antibodies can be used in methods known in the art relating to the localization and activity of the lats protein sequences of the invention, e.g., for imaging these proteins, measuring levels thereof in appropriate physiological samples, in diagnostic methods,  
25 etc.

In another embodiment of the invention (see *infra*), anti-lats antibodies and fragments thereof containing the binding domain are Therapeutics.

### 30 5.6. LATS PROTEINS, DERIVATIVES AND ANALOGS

The invention further relates to lats proteins, and derivatives (including but not limited to fragments) and analogs of lats proteins. Nucleic acids encoding lats protein derivatives and protein analogs are also provided.  
35 In one embodiment, the lats proteins are encoded by the lats nucleic acids described in Section 5.1 *supra*. In particular aspects, the proteins, derivatives, or analogs are of lats

pr teins of animals, .g., fly, frog, mouse, rat, pig, cow, dog, m nkey, human, or of plants.

The production and use of derivatives and analogs related to lats are within the scope of the present invention. In a specific embodiment, the derivative or analog is functionally active, i.e., capable of exhibiting one or more functional activities associated with a full-length, wild-type lats protein. As one example, such derivatives or analogs which have the desired immunogenicity or antigenicity can be used, for example, in immunoassays, for immunization, for inhibition of lats activity, etc. As another example, such derivatives or analogs which have the desired kinase activity, or which are phosphorylated or dephosphorylated, are provided. Derivatives or analogs that retain, or alternatively lack or inhibit, a desired lats property of interest (e.g., binding to an SH3-domain-containing protein or other lats binding partner, kinase activity, inhibition of cell proliferation, tumor inhibition), can be used as inducers, or inhibitors, respectively, of such property and its physiological correlates. A specific embodiment relates to a lats fragment that can be bound by an anti-lats antibody. Derivatives or analogs of lats can be tested for the desired activity by procedures known in the art, including but not limited to the assays described in Sections 5.7 and 5.9.

In particular, lats derivatives can be made by altering lats sequences by substitutions, additions or deletions that provide for functionally equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence as a lats gene may be used in the practice of the present invention. These include but are not limited to nucleotide sequences comprising all or portions of lats genes which are altered by the substitution of different codons that encode a functionally equivalent amino acid residue within the sequence, thus producing a silent change. Likewise, the lats derivatives of the invention include, but



are not limited to, those containing, as a primary amino acid sequence, all or part of the amino acid sequence of a lats protein including altered sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence resulting in a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid.

In a specific embodiment of the invention, proteins consisting of or comprising a fragment of a lats protein consisting of at least 10 (continuous) amino acids of the lats protein is provided. In other embodiments, the fragment consists of at least 20 or 50 amino acids of the lats protein. In specific embodiments, such fragments are not larger than 35, 100 or 200 amino acids. Derivatives or analogs of lats include but are not limited to those molecules comprising regions that are substantially homologous to lats or fragments thereof (e.g., in various embodiments, at least 60% or 70% or 80% or 90% or 95% identity over an amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art) or whose encoding nucleic acid is capable of hybridizing to a coding lats sequence, under stringent, moderately stringent, or nonstringent conditions.

Th lats d rivatives and analogs of the inv ntion can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level. For example, the cloned lats gene  
5 sequence can be modified by any of numerous strategies known in the art (Maniatis, T., 1990, Molecular Cloning, A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). The sequence can be cleaved at appropriate sites with restriction endonuclease(s),  
10 followed by further enzymatic modification if desired, isolated, and ligated in vitro. In the production of the gene encoding a derivative or analog of lats, care should be taken to ensure that the modified gene remains within the same translational reading frame as lats, uninterrupted by  
15 translational stop signals, in the gene region where the desired lats activity is encoded.

Additionally, the lats-encoding nucleic acid sequence can be mutated in vitro or in vivo, to create and/or destroy translation, initiation, and/or termination  
20 sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further in vitro modification. Any technique for mutagenesis known in the art can be used, including but not limited to, chemical  
25 mutagenesis, in vitro site-directed mutagenesis (Hutchinson, C., et al., 1978, J. Biol. Chem 253:6551), use of TAB® linkers (Pharmacia), etc.

Manipulations of the lats sequence may also be made at the protein level. Included within the scope of the  
30 invention are lats protein fragments or other derivatives or analogs which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to  
35 an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited to specific chemical

cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease,  $\text{NaBH}_4$ ; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

- 5 In addition, analogs and derivatives of lats can be chemically synthesized. For example, a peptide corresponding to a portion of a lats protein which comprises the desired domain (see Section 5.6.1), or which mediates the desired activity *in vitro*, can be synthesized by use of a peptide
- 10 synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the lats sequence. Non-classical amino acids include but are not limited to the D-isomers of the common amino acids,  $\alpha$ -amino isobutyric acid,
- 15 4-aminobutyric acid, Abu, 2-amino butyric acid,  $\gamma$ -Abu,  $\epsilon$ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine,
- 20 cyclohexylalanine,  $\beta$ -alanine, fluoro-amino acids, designer amino acids such as  $\beta$ -methyl amino acids,  $\text{C}\alpha$ -methyl amino acids,  $\text{N}\alpha$ -methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).
- 25 In a specific embodiment, the lats derivative is a chimeric, or fusion, protein comprising a lats protein or fragment thereof (preferably consisting of at least a domain or motif of the lats protein, or at least 10 amino acids of the lats protein) joined at its amino- or carboxy-terminus
- 30 via a peptide bond to an amino acid sequence of a different protein. In one embodiment, such a chimeric protein is produced by recombinant expression of a nucleic acid encoding the protein (comprising a lats-coding sequence joined in-frame to a coding sequence for a different protein). Such a
- 35 chimeric product can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other by methods known in the art, in the

proper coding frame, and expressing the chimeric product by methods commonly known in the art. Alternatively, such a chimeric product may be made by protein synthetic techniques, e.g., by use of a peptide synthesizer. Chimeric genes comprising portions of lats fused to any heterologous protein-encoding sequences may be constructed. A specific embodiment relates to a chimeric protein comprising a fragment of lats of at least six amino acids.

In another specific embodiment, the lats derivative is a molecule comprising a region of homology with a lats protein. By way of example, in various embodiments, a first protein region can be considered "homologous" to a second protein region when the amino acid sequence of the first region is at least 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, or 95% identical, when compared to any sequence in the second region of an equal number of amino acids as the number contained in the first region or when compared to an aligned sequence of the second region that has been aligned by a computer homology program known in the art. For example, a molecule can comprise one or more regions homologous to a lats domain (see Section 5.6.1) or a portion thereof.

Other specific embodiments of derivatives and analogs are described in the subsection below and examples sections *infra*.

25

#### 5.6.1. DERIVATIVES OF LATS CONTAINING ONE OR MORE DOMAINS OF THE PROTEIN

In a specific embodiment, the invention relates to lats derivatives and analogs, in particular lats fragments and derivatives of such fragments, that comprise, or alternatively consist of, one or more domains of a lats protein, including but not limited to a lats C-terminal domain 3 (LCD3), lats C-terminal domain 2 (LCD2), lats C-terminal domain 1 (LCD1), kinase domain, kinase subdomains, lats flanking domain (LFD) (amino-terminal to the kinase domain), lats split domain 1 (LSD1), lats split domain 2 (LSD2), SH3-binding domain, and opa repeat domain, functional

(e.g., binding) fragments of any of the foregoing, or any combination of the foregoing. In particular examples relating to the human, mouse and *Drosophila* lats proteins, such domains are identified in Examples Sections 6 and 8, and 5 in Figures 6A, 6B, and 13.

A specific embodiment relates to molecules comprising specific fragments of lats that are those fragments in the respective lats protein most homologous to specific fragments of a human or mouse lats protein. A  
10 fragment comprising a domain of a lats homolog can be identified by protein analysis methods as described in Sections 5.3.2 or 6.

In a specific embodiment, a lats protein, derivative or analog is provided that has a kinase domain and  
15 has a phosphorylated serine situated within 20 residues upstream of an Ala-Pro-Glu consensus in subdomain eight of its kinase domain. In another embodiment, a lats protein derivative or analog is provided with a kinase domain and with a dephosphorylated serine situated within 20 residues  
20 upstream of an Ala-Pro-Glu consensus in subdomain eight of its kinase domain, or in which the serine situated within 20 residues upstream of that consensus has been deleted or substituted by another amino acid. In a specific embodiment, the invention provides various phosphorylated and  
25 dephosphorylated forms of the lats protein, derivative, or analog that are active kinase forms. Both phosphorylation and dephosphorylation of lats at different residues could potentially activate lats. In another specific embodiment, the invention provides various phosphorylated and  
30 dephosphorylated forms of the lats protein, derivative or analog that are inactive kinase forms. Phosphorylation can be carried out by any methods known in the art, e.g., by use of a kinase. Dephosphorylation can be carried out by use of any methods known in the art, e.g., by use of a phosphatase.  
35 Another specific embodiment relates to a derivative or analog of a lats protein that is a dominant-active protein kinase. Such a derivative or analog comprises a lats kinase

domain that has been mutated so as to be dominantly active ( exhibit constitutiv ly active kinase activity). It is known that acidic residues such as Glu and Asp sometimes mimic a phosphorylated residue, and changing the phosphorylatable Ser 5 or Thr residue in subdomain eight into a Glu or Asp residue has been previously used to produce constitutively active kinases (Mansour et al., 1994, Science 265:966-970). Thus, changing a serine or threonine residue situated within 20 residues upstream of an Ala-Pro-Glu consensus in subdomain 10 eight of a lats kinase domain into another residue (e.g., Glu, Asp) may be used to make a dominant-active lats protein kinase. For example, changing Ser914 in *Drosophila* lats, or changing Ser909 in h-lats, into a Glu residue could produce a dominant active lats kinase.

15 Another specific embodiment relates to a derivative or analog of lats that is a dominant-negative protein kinase. Protein kinases can be mutated into dominant negative forms. Expression of a dominant negative protein kinase can suppress the activity of the wild-type form of the same kinase.

20 Dominant negative forms of protein kinases are often obtained by expressing an inactive form of a kinase (Milarski and Saltiel, 1994, J. Biol. Chem. 269(33):21239-21243) or by expressing a noncatalytic domain of a kinase (Lu and Means, 1994, EMBO J. 12:2103-2113; Yarden et al., 1992, EMBO J. 25 11:2159-2166). Thus, a lats dominant-negative kinase can be obtained by mutating the kinase domain so as to be inactive (e.g., by deletion and/or point mutation). By way of example, a lats derivative that is a dominant-negative kinase is a lats protein that lacks a kinase domain but comprises 30 one or more of the other domains of the lats protein; e.g., a lats protein derivative truncated at about the beginning of the kinase domain (i.e., a lats fragment containing only sequences amino-terminal to the kinase domain). By way of another example, a lats derivative that is a dominant-

35 negative kinase is a lats protein in which one of the residues conserved among s rine/threonine kinases (see Hanks

et al., 1988, Science 241:42-52) is mutated (deleted or substituted by a different residue).

In another specific embodiment, a molecule is provided that comprises one or more domains (or functional portion thereof) of a lats protein but that also lacks one or more domains (or functional portion thereof) of a lats protein. In particular examples, lats protein derivatives are provided that lack an opa repeat domain. By way of another example, such a protein may also lack all or a portion of the kinase domain, but retain at least the SH3-binding domain of a lats protein. In another embodiment, a molecule is provided that comprises one or more domains (or functional portion thereof) of a lats protein, and that has one or more mutant (e.g., due to deletion or point mutation(s)) domains of a lats protein (e.g., such that the mutant domain has decreased function). By way of example, the kinase domain may be mutant so as to have reduced, absent, or increased kinase activity.

#### 20 5.7. ASSAYS OF LATS PROTEINS, DERIVATIVES AND ANALOGS

The functional activity of lats proteins, derivatives and analogs can be assayed by various methods.

For example, in one embodiment, where one is assaying for the ability to bind or compete with wild-type lats for binding to anti-lats antibody, various immunoassays known in the art can be used, including but not limited to competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (e.g., gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In.

on embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labelled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention.

In another embodiment, where a lats-binding protein is identified, the binding can be assayed, e.g., by means well-known in the art. In another embodiment, physiological correlates of lats binding to its substrates (signal transduction) can be assayed.

In another embodiment, kinase assays can be used to measure lats kinase activity. Such assays can be carried out by methods well known in the art. By way of example, a lats protein is contacted with a substrate (e.g., a known substrate of serine/threonine kinases) in the presence of a  $^{32}\text{P}$ -labeled phosphate donor, and any phosphorylation of the substrate is detected or measured.

In another embodiment, in insect or other model systems, genetic studies can be done to study the phenotypic effect of a lats mutant that is a derivative or analog of wild-type lats (see Section 6, *infra*).

In addition, assays that can be used to detect or measure the ability to inhibit, or alternatively promote, cell proliferation are described in Section 5.9.

Other methods will be known to the skilled artisan and are within the scope of the invention.

#### 5.8. THERAPEUTIC USES

The invention provides for treatment or prevention of various diseases and disorders by administration of a therapeutic compound (termed herein "Therapeutic"). Such "Therapeutics" include but are not limited to: lats proteins and analogs and derivatives (including fragments) thereof (e.g., as described hereinabove); antibodies thereto (as



described hereinabove); nucleic acids encoding the lats  
prot ins, analogs, or derivatives ( .g., as d scribed  
h r inabove); lats antisense nucleic acids, and lats agonists  
and antagonists. Disorders involving c ll overproliferation  
5 are treated or prevented by administration of a Therapeutic  
that promotes lats function. Disorders in which cell  
proliferation is deficient or is desired are treated or  
prevented by administration of a Therapeutic that antagonizes  
(inhibits) lats function. The above is described in detail  
10 in the subsections below.

Generally, administration of products of a species  
origin or species reactivity (in the case of antibodies) that  
is the same species as that of the patient is preferred.  
Thus, in a preferred embodiment, a human lats protein,  
15 derivative, or analog, or nucleic acid, or an antibody to a  
human lats protein, is therapeutically or prophylactically  
administered to a human patient.

Additional descriptions and sources of Therapeutics  
that can be used according to the invention are found in  
20 Sections 5.1 through 5.7 herein.

#### 5.8.1. TREATMENT AND PREVENTION OF DISORDERS INVOLVING OVERPROLIFERATION OF CELLS

Diseases and disorders involving cell  
25 overproliferation are treated or prevented by administration  
of a Therapeutic that promotes (i.e., increases or supplies)  
lats function. Examples of such a Therapeutic include but  
are not limited to lats proteins, derivatives, or fragments  
that are functionally active, particularly that are active in  
30 inhibiting cell proliferation (e.g., as demonstrated in in  
vitro assays or in animal models or in *Drosophila*), and  
nucleic acids encoding a lats protein or functionally active  
derivative or fragment thereof (e.g., for use in gene  
therapy). Other Therapeutics that can be used, e.g., lats  
35 agonists, can be identified using in vitro assays or animal  
models, or assays in *Drosophila*, examples of which are  
described *infra*.

In specific embodiments, Therapeutics that promote lats function are administered therapeutically (including prophylactically): (1) in diseases or disorders involving an absence or decreased (relative to normal or desired) level of lats protein or function, for example, in patients where lats protein is lacking, genetically defective, biologically inactive or underactive, or underexpressed; or (2) in diseases or disorders wherein *in vitro* (or *in vivo*) assays (see *infra*) indicate the utility of lats agonist administration. The absence or decreased level in lats protein or function can be readily detected, e.g., by obtaining a patient tissue sample (e.g., from biopsy tissue) and assaying it *in vitro* for RNA or protein levels, structure and/or activity of the expressed lats RNA or protein. Many methods standard in the art can be thus employed, including but not limited to kinase assays, immunoassays to detect and/or visualize lats protein (e.g., Western blot, immunoprecipitation followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect lats expression by detecting and/or visualizing lats mRNA (e.g., Northern assays, dot blots, *in situ* hybridization, etc.), etc.

Diseases and disorders involving cell overproliferation that can be treated or prevented include but are not limited to malignancies, premalignant conditions (e.g., hyperplasia, metaplasia, dysplasia), benign tumors, hyperproliferative disorders, benign dysproliferative disorders, etc. Examples of these are detailed below.

In a specific embodiment, the Therapeutic used, that promotes lats function, is a lats protein, derivative or analog comprising a lats kinase domain (and optionally also a lats LFD, or the remainder of the lats sequence) in which a serine within 20 residues upstream of the Ala-Pro-Glu consensus in subdomain eight of the kinase domain is phosphorylated or substituted by another residue (e.g., Glu, Asp).

In another specific embodiment, the Therapeutic used, that promotes lats function, is a derivative or analog comprising a kinase domain of a lats protein that has been mutated so as to be dominantly active.

5

#### 5.8.1.1. MALIGNANCIES

Malignancies and related disorders that can be treated or prevented by administration of a Therapeutic that promotes lats function include but are not limited to those  
 10 listed in Table 1 (for a review of such disorders, see Fishman et al., 1985, *Medicine*, 2d Ed., J.B. Lippincott Co., Philadelphia):

15

TABLE 1  
MALIGNANCIES AND RELATED DISORDERS

	Leukemia
	acute leukemia
	acute lymphocytic leukemia
20	acute myelocytic leukemia
	myeloblastic
	promyelocytic
	myelomonocytic
	monocytic
	erythroleukemia
	chronic leukemia
	chronic myelocytic (granulocytic) leukemia
25	chronic lymphocytic leukemia
	Polycythemia vera
	Lymphoma
	Hodgkin's disease
	non-Hodgkin's disease
	Multiple myeloma
	Waldenström's macroglobulinemia
	Heavy chain disease
30	Solid tumors
	sarcomas and carcinomas
	fibrosarcoma
	myxosarcoma
	liposarcoma
	chondrosarcoma
	osteogenic sarcoma
35	chordoma
	angiosarcoma
	endotheliosarcoma
	lymphangiosarcoma

	lymphangioendoth liosarcoma
	synovioma
	mesothelioma
	Ewing's tumor
	l iomyosarcoma
5	rhabdomyosarcoma
	colon carcinoma
	pancreatic cancer
	breast cancer
	ovarian cancer
	prostate cancer
	squamous cell carcinoma
	basal cell carcinoma
10	adenocarcinoma
	sweat gland carcinoma
	sebaceous gland carcinoma
	papillary carcinoma
	papillary adenocarcinomas
	cystadenocarcinoma
	medullary carcinoma
	bronchogenic carcinoma
15	renal cell carcinoma
	hepatoma
	bile duct carcinoma
	choriocarcinoma
	seminoma
	embryonal carcinoma
	Wilms' tumor
20	cervical cancer
	uterine cancer
	testicular tumor
	lung carcinoma
	small cell lung carcinoma
	bladder carcinoma
	epithelial carcinoma
	glioma
25	astrocytoma
	medulloblastoma
	craniopharyngioma
	ependymoma
	pinealoma
	hemangioblastoma
	acoustic neuroma
30	oligodendroglioma
	menangioma
	melanoma
	neuroblastoma
	retinoblastoma

---

35            In specific embodiments, malignancy or  
dysproliferative changes (such as metaplasias and

dysplasias), or hyperproliferative disorders, are treated or prevented in the bladder, breast, colon, lung, melanoma, pancreas, or uterus. In other specific embodiments, sarcoma, or leukemia is treated or prevented.

5

#### 5.8.1.2. PREMALIGNANT CONDITIONS

The Therapeutics of the invention that promote lats activity can also be administered to treat premalignant conditions and to prevent progression to a neoplastic or malignant state, including but not limited to those disorders listed in Table 1. Such prophylactic or therapeutic use is indicated in conditions known or suspected of preceding progression to neoplasia or cancer, in particular, where non-neoplastic cell growth consisting of hyperplasia, metaplasia, or most particularly, dysplasia has occurred (for review of such abnormal growth conditions, see Robbins and Angell, 1976, *Basic Pathology*, 2d Ed., W.B. Saunders Co., Philadelphia, pp. 68-79.) Hyperplasia is a form of controlled cell proliferation involving an increase in cell number in a tissue or organ, without significant alteration in structure or function. As but one example, endometrial hyperplasia often precedes endometrial cancer. Metaplasia is a form of controlled cell growth in which one type of adult or fully differentiated cell substitutes for another type of adult cell. Metaplasia can occur in epithelial or connective tissue cells. Atypical metaplasia involves a somewhat disorderly metaplastic epithelium. Dysplasia is frequently a forerunner of cancer, and is found mainly in the epithelia; it is the most disorderly form of non-neoplastic cell growth, involving a loss in individual cell uniformity and in the architectural orientation of cells. Dysplastic cells often have abnormally large, deeply stained nuclei, and exhibit pleomorphism. Dysplasia characteristically occurs where there exists chronic irritation or inflammation, and is often found in the cervix, respiratory passages, oral cavity, and gall bladder.

Alternatively or in addition to the presence of abnormal cell growth characterized as hyperplasia, metaplasia, or dysplasia, the presence of one or more characteristics of a transformed phenotype, or of a malignant phenotype, displayed *in vivo* or displayed *in vitro* by a cell sample from a patient, can indicate the desirability of prophylactic/therapeutic administration of a Therapeutic that promotes lats function. As mentioned *supra*, such characteristics of a transformed phenotype include morphology changes, looser substratum attachment, loss of contact inhibition, loss of anchorage dependence, protease release, increased sugar transport, decreased serum requirement, expression of fetal antigens, disappearance of the 250,000 dalton cell surface protein, etc. (see also *id.*, at pp. 84-90 for characteristics associated with a transformed or malignant phenotype).

In a specific embodiment, leukoplakia, a benign-appearing hyperplastic or dysplastic lesion of the epithelium, or Bowen's disease, a carcinoma *in situ*, are pre-neoplastic lesions indicative of the desirability of prophylactic intervention.

In another embodiment, fibrocystic disease (cystic hyperplasia, mammary dysplasia, particularly adenosis (benign epithelial hyperplasia)) is indicative of the desirability of prophylactic intervention.

In other embodiments, a patient which exhibits one or more of the following predisposing factors for malignancy is treated by administration of an effective amount of a Therapeutic: a chromosomal translocation associated with a malignancy (e.g., the Philadelphia chromosome for chronic myelogenous leukemia, t(14;18) for follicular lymphoma, etc.), familial polyposis or Gardner's syndrome (possible forerunners of colon cancer), benign monoclonal gammopathy (a possible forerunner of multiple myeloma), and a first degree kinship with persons having a cancer or precancerous disease showing a Mendelian (genetic) inheritance pattern (e.g., familial polyposis of the colon, Gardner's syndrome,

hereditary xostosis, poly ndocrine ad nomatosis, medullary thyroid carcinoma with amyloid production and phe chromocytoma, Peutz-Jeghers syndrome, neurofibromatosis of Von Recklinghausen, retinoblastoma, carotid body tumor, 5 cutaneous melanocarcinoma, intraocular melanocarcinoma, xeroderma pigmentosum, ataxia telangiectasia, Chediak-Higashi syndrome, albinism, Fanconi's aplastic anemia, and Bloom's syndrome; see Robbins and Angell, 1976, *Basic Pathology*, 2d Ed., W.B. Saunders Co., Philadelphia, pp. 112-113) etc.)

10 In another specific embodiment, a Therapeutic of the invention is administered to a human patient to prevent progression to breast, colon, lung, pancreatic, or uterine cancer, or melanoma or sarcoma.

15 5.8.1.3. HYPERPROLIFERATIVE AND DYSPROLIFERATIVE DISORDERS

In another embodiment of the invention, a Therapeutic that promotes lats activity is used to treat or prevent hyperproliferative or benign dysproliferative 20 disorders. Specific embodiments are directed to treatment or prevention of cirrhosis of the liver (a condition in which scarring has overtaken normal liver regeneration processes), treatment of keloid (hypertrophic scar) formation (disfiguring of the skin in which the scarring process 25 interferes with normal renewal), psoriasis (a common skin condition characterized by excessive proliferation of the skin and delay in proper cell fate determination), benign tumors, fibrocystic conditions, and tissue hypertrophy (e.g., prostatic hyperplasia).

30 5.8.1.4. GENE THERAPY

In a specific embodiment, nucleic acids comprising a sequence encoding a lats protein or functional derivative thereof, are administered to promote lats function, by way of 35 gene therapy. Gene therapy refers to therapy performed by the administration of a nucleic acid to a subject. In this embodiment of the invention, the nucleic acid produces its

encoded prot in that mediates a therapeutic effect by promoting lats function.

Any of the m thods for gene therapy available in the art can be used according to the present invention.

5 Exemplary methods are described below.

For general reviews of the methods of gene therapy, see Goldspiel et al., 1993, Clinical Pharmacy 12:488-505; Wu and Wu, 1991, Biotherapy 3:87-95; Tolstoshev, 1993, Ann. Rev. Pharmacol. Toxicol. 32:573-596; Mulligan, 1993, Science  
10 260:926-932; and Morgan and Anderson, 1993, Ann. Rev. Biochem. 62:191-217; May, 1993, TIBTECH 11(5):155-215). Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), 1993, Current Protocols in Molecular Biology, John  
15 Wiley & Sons, NY; and Kriegler, 1990, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY.

In a preferred aspect, the Therapeutic comprises a lats nucleic acid that is part of an expression vector that expresses a lats protein or fragment or chimeric protein  
20 thereof in a suitable host. In particular, such a nucleic acid has a promoter operably linked to the lats coding region, said promoter being inducible or constitutive, and, optionally, tissue-specific. In another particular embodiment, a nucleic acid molecule is used in which the lats  
25 coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the lats nucleic acid (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et  
30 al., 1989, Nature 342:435-438).

Delivery of the nucleic acid into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, cells are first transformed with the  
35 nucleic acid in vitro, then transplanted into the patient. These two approaches are known, respectively, as in vivo or ex vivo gen therapy.



In a specific embodiment, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, e.g., by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by infection using a defective or attenuated retroviral or other viral vector (see U.S. Patent No. 4,980,286), or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering it in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432) (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, a nucleic acid-ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo* for cell specific uptake and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180 dated April 16, 1992 (Wu et al.); WO 92/22635 dated December 23, 1992 (Wilson et al.); WO92/20316 dated November 26, 1992 (Findeis et al.); WO93/14188 dated July 22, 1993 (Clarke et al.), WO 93/20221 dated October 14, 1993 (Young)). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438).

In a specific embodiment, a viral vector that contains the *lats* nucleic acid is used. For example, a retroviral vector can be used (see Miller et al., 1993, Meth. Enzymol. 217:581-599). These retroviral vectors have been

modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA. The late nucleic acid to be used in gene therapy is cloned into the vector, which facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., 1994, *Biotherapy* 6:291-302, which describes the use of a retroviral vector to deliver the *mdr1* gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are: Clowes et al., 1994, *J. Clin. Invest.* 93:644-651; Kiem et al., 1994, *Blood* 83:1467-1473; Salmons and Gunzberg, 1993, *Human Gene Therapy* 4:129-141; and Grossman and Wilson, 1993, *Curr. Opin. in Genetics and Devel.* 3:110-114.

Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Kozarsky and Wilson, 1993, *Current Opinion in Genetics and Development* 3:499-503 present a review of adenovirus-based gene therapy. Bout et al., 1994, *Human Gene Therapy* 5:3-10, demonstrated the use of adenovirus vectors to transfer genes to the respiratory epithelia of rhesus monkeys. Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., 1991, *Science* 252:431-434; Rosenfeld et al., 1992, *Cell* 68:143-155; and Mastrangeli et al., 1993, *J. Clin. Invest.* 91:225-234.

Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., 1993, *Proc. Soc. Exp. Biol. Med.* 204:289-300).

Another approach to gene therapy involves transferring a gene to cells in tissue culture by such

methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

In this embodiment, the nucleic acid is introduced into a cell prior to administration *in vivo* of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see e.g., Loeffler and Behr, 1993, Meth. Enzymol. 217:599-618; Cohen et al., 1993, Meth. Enzymol. 217:618-644; Cline, 1985, Pharmac. Ther. 29:69-92) and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

The resulting recombinant cells can be delivered to a patient by various methods known in the art. In a preferred embodiment, epithelial cells are injected, e.g., subcutaneously. In another embodiment, recombinant skin cells may be applied as a skin graft onto the patient. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art.

Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as T lymphocytes, B lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

In an embodiment in which recombinant cells are used in gene therapy, a nucleic acid is introduced into the cells such that it is expressible by the cells or their progeny, and the recombinant cells are then administered *in vivo* for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem and/or progenitor cells which can be isolated and maintained *in vitro* can potentially be used in accordance with this embodiment of the present invention. Such stem cells include but are not limited to hematopoietic stem cells (HSC), stem cells of epithelial tissues such as the skin and the lining of the gut, embryonic heart muscle cells, liver stem cells (PCT Publication WO 94/08598, dated April 28, 1994), and neural stem cells (Stemple and Anderson, 1992, *Cell* 71:973-985).

Epithelial stem cells (ESCs) or keratinocytes can be obtained from tissues such as the skin and the lining of the gut by known procedures (Rheinwald, 1980, *Meth. Cell Bio.* 21A:229). In stratified epithelial tissue such as the skin, renewal occurs by mitosis of stem cells within the germinal layer, the layer closest to the basal lamina. Stem cells within the lining of the gut provide for a rapid renewal rate of this tissue. ESCs or keratinocytes obtained from the skin or lining of the gut of a patient or donor can be grown in tissue culture (Rheinwald, 1980, *Meth. Cell Bio.* 21A:229; Pittelkow and Scott, 1986, *Mayo Clinic Proc.* 61:771). If the

ESCs are provided by a donor, a method for suppression of host versus graft reaction activity (e.g., irradiation, drug or antibody administration to promote moderate immunosuppression) can also be used.

5           With respect to hematopoietic stem cells (HSC), any technique which provides for the isolation, propagation, and maintenance *in vitro* of HSC can be used in this embodiment of the invention. Techniques by which this may be accomplished include (a) the isolation and establishment of HSC cultures  
10 from bone marrow cells isolated from the future host, or a donor, or (b) the use of previously established long-term HSC cultures, which may be allogeneic or xenogeneic. Non-autologous HSC are used preferably in conjunction with a method of suppressing transplantation immune reactions of the  
15 future host/patient. In a particular embodiment of the present invention, human bone marrow cells can be obtained from the posterior iliac crest by needle aspiration (see, e.g., Kodo et al., 1984, J. Clin. Invest. 73:1377-1384). In a preferred embodiment of the present invention, the HSCs can  
20 be made highly enriched or in substantially pure form. This enrichment can be accomplished before, during, or after long-term culturing, and can be done by any techniques known in the art. Long-term cultures of bone marrow cells can be established and maintained by using, for example, modified  
25 Dexter cell culture techniques (Dexter et al., 1977, J. Cell Physiol. 91:335) or Witlock-Witte culture techniques (Witlock and Witte, 1982, Proc. Natl. Acad. Sci. USA 79:3608-3612).

          In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an  
30 inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription.

          Additional methods that can be adapted for use to  
35 deliver a nucleic acid encoding a protein or functional derivative thereof are described in Section 5.8.2.2.2.

### 5.8.2. TREATMENT AND PREVENTION OF DISORDERS IN WHICH CELL PROLIFERATION IS DESIRED

Diseases and disorders involving a deficiency in cell proliferation (growth) or in which cell proliferation is otherwise desirable for treatment or prevention, are treated or prevented by administration of a Therapeutic that antagonizes (inhibits) *lats* function (in particular, *lats*-mediated inhibition of cell proliferation). Therapeutics that can be used include but are not limited to anti-*lats* antibodies (and fragments and derivatives thereof containing the binding region thereof), *lats* derivatives or analogs that are dominant-negative kinases, *lats* antisense nucleic acids, and *lats* nucleic acids that are dysfunctional (e.g., due to a heterologous (non-*lats* sequence) insertion within the *lats* coding sequence) that are used to "knockout" endogenous *lats* function by homologous recombination (see, e.g., Capecchi, 1989, Science 244:1288-1292). In a specific embodiment of the invention, a nucleic acid containing a portion of a *lats* gene in which *lats* sequences flank (are both 5' and 3' to) a different gene sequence, is used, as a *lats* antagonist, to promote *lats* inactivation by homologous recombination (see also Koller and Smithies, 1989, Proc. Natl. Acad. Sci. USA 86:8932-8935; Zijlstra et al., 1989, Nature 342:435-438). Other Therapeutics that inhibit *lats* function can be identified by use of known convenient *in vitro* assays, e.g., based on their ability to inhibit binding of *lats* to another protein (e.g., an SH3-domain containing protein), or inhibit any known *lats* function, as preferably assayed *in vitro* or in cell culture, although genetic assays (e.g., in *Drosophila*) may also be employed. Preferably, suitable *in vitro* or *in vivo* assays, are utilized to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In specific embodiments, Therapeutics that inhibit *lats* function are administered therapeutically (including prophylactically): (1) in diseases or disorders involving an increased (relative to normal or desired) level of *lats*

protein or function, for example, in patients where lats  
protein is overactive or overexpressed; or (2) in diseases or  
disorders wherein *in vitro* (or *in vivo*) assays (see *infra*)  
indicate the utility of lats antagonist administration. The  
5 increased levels in lats protein or function can be readily  
detected, e.g., by quantifying protein and/or RNA, by  
obtaining a patient tissue sample (e.g., from biopsy tissue)  
and assaying it *in vitro* for RNA or protein levels, structure  
and/or activity of the expressed lats RNA or protein. Many  
10 methods standard in the art can be thus employed, including  
but not limited to kinase assays, immunoassays to detect  
and/or visualize lats protein (e.g., Western blot,  
immunoprecipitation followed by sodium dodecyl sulfate  
polyacrylamide gel electrophoresis, immunocytochemistry,  
15 etc.) and/or hybridization assays to detect lats expression  
by detecting and/or visualizing respectively lats mRNA (e.g.,  
Northern assays, dot blots, *in situ* hybridization, etc.),  
etc.

Diseases and disorders involving a deficiency in  
20 cell proliferation or in which cell proliferation is desired  
for treatment or prevention, and that can be treated or  
prevented by inhibiting lats function, include but are not  
limited to degenerative disorders, growth deficiencies,  
hypoproliferative disorders, physical trauma, lesions, and  
25 wounds; for example, to promote wound healing, or to promote  
regeneration in degenerated, lesioned or injured tissues,  
etc. In a specific embodiment, nervous system disorders are  
treated. In another specific embodiment, a disorder that is  
not of the nervous system is treated.

30 Lesions which may be treated according to the  
present invention include but are not limited to the  
following lesions:

- (i) traumatic lesions, including lesions caused by  
physical injury or associated with surgery;
- 35 (ii) ischemic lesions, in which a lack of oxygen  
results in cell injury or death, e.g.,

- myocardial or cerebral infarction or ischemia,  
or spinal cord infarction or ischemia;
- 5 (iii) malignant lesions, in which cells are  
destroyed or injured by malignant tissue;
- (iv) 10 infectious lesions, in which tissue is  
destroyed or injured as a result of infection,  
for example, by an abscess or associated with  
infection by human immunodeficiency virus,  
herpes zoster, or herpes simplex virus or with  
15 Lyme disease, tuberculosis, syphilis;
- (v) degenerative lesions, in which tissue is  
destroyed or injured as a result of a  
degenerative process, including but not  
limited to nervous system degeneration  
associated with Parkinson's disease,  
Alzheimer's disease, Huntington's chorea, or  
amyotrophic lateral sclerosis;
- 20 (vi) lesions associated with nutritional diseases  
or disorders, in which tissue is destroyed or  
injured by a nutritional disorder or disorder  
of metabolism including but not limited to,  
vitamin B12 deficiency, folic acid deficiency,  
Wernicke disease, tobacco-alcohol amblyopia,  
Marchiafava-Bignami disease (primary  
25 degeneration of the corpus callosum), and  
alcoholic cerebellar degeneration;
- (vii) lesions associated with systemic diseases  
including but not limited to diabetes or  
systemic lupus erythematosus;
- 30 (viii) lesions caused by toxic substances including  
alcohol, lead, or other toxins; and
- (ix) demyelinated lesions of the nervous system, in  
which a portion of the nervous system is  
destroyed or injured by a demyelinating  
35 disease including but not limited to multiple  
sclerosis, human immunodeficiency virus-  
associated myelopathy, transverse myelopathy



or various tiologi s, progressive multifocal  
l ukoencephalopathy, and central pontine  
my linolysis.

Nervous system lesions which may be treated in a  
5 patient (including human and non-human mammalian patients)  
according to the invention include but are not limited to the  
lesions of either the central (including spinal cord, brain)  
or peripheral nervous systems.

Therapeutics which are useful according to this  
10 embodiment of the invention for treatment of a disorder may  
be selected by testing for biological activity in promoting  
the survival or differentiation of cells (see also Section  
5.9). For example, in a specific embodiment relating to  
therapy of the nervous system, a Therapeutic which elicits  
15 one of the following effects may be useful according to the  
invention:

(i) increased sprouting of neurons in culture or  
*in vivo*;

(ii) increased production of a neuron-associated  
20 molecule in culture or *in vivo*, e.g., cholin  
acetyltransferase or acetylcholinesterase with  
respect to motor neurons; or

(iii) decreased symptoms of neuron dysfunction in  
*vivo*.

25 Such effects may be measured by any method known in the art.  
In preferred, non-limiting embodiments, increased sprouting  
of neurons may be detected by methods set forth in Pestronk  
et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981,  
Ann. Rev. Neurosci. 4:17-42); and increased production of  
30 neuron-associated molecules may be measured by bioassay,  
enzymatic assay, antibody binding, Northern blot assay, etc.,  
depending on the molecule to be measured.

#### 35 5.8.2.1. ANTISENSE REGULATION OF LATS EXPRESSION

In a specific embodiment, lats function is  
inhibit d by us of lats antisense nucleic acids. The

present invention provides the therapeutic or prophylactic use of nucleic acids of at least six nucleotides that are antisense to a gene or cDNA encoding lats or a portion thereof. A lats "antisense" nucleic acid as used herein  
5 refers to a nucleic acid capable of hybridizing to a portion of a lats RNA (preferably mRNA) by virtue of some sequence complementarity. The antisense nucleic acid may be complementary to a coding and/or noncoding region of a lats mRNA. Such antisense nucleic acids have utility as  
10 Therapeutics that inhibits lats function, and can be used in the treatment or prevention of disorders as described *supra* in Section 5.8.2 and its subsections.

The antisense nucleic acids of the invention can be oligonucleotides that are double-stranded or single-stranded,  
15 RNA or DNA or a modification or derivative thereof, which can be directly administered to a cell, or which can be produced intracellularly by transcription of exogenous, introduced sequences.

In a specific embodiment, the lats antisense  
20 nucleic acids provided by the instant invention can be used to promote regeneration or wound healing or to promote growth (larger size).

The invention further provides pharmaceutical compositions comprising an effective amount of the lats  
25 antisense nucleic acids of the invention in a pharmaceutically acceptable carrier, as described *infra*.

In another embodiment, the invention is directed to methods for inhibiting the expression of a lats nucleic acid sequence in a prokaryotic or eukaryotic cell comprising  
30 providing the cell with an effective amount of a composition comprising an lats antisense nucleic acid of the invention.

Lats antisense nucleic acids and their uses are described in detail below.

#### 5.8.2.1.1. LATS ANTISENSE NUCLEIC ACIDS

The lats antisense nucleic acids are of at least six nucleotides and are preferably oligonucleotides (ranging from 6 to about 50 oligonucleotides). In specific aspects, 5 the oligonucleotide is at least 10 nucleotides, at least 15 nucleotides, at least 100 nucleotides, or at least 200 nucleotides. The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The 10 oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone. The oligonucleotide may include other appending groups such as peptides, or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 15 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. WO 88/09810, published December 15, 1988) or blood-brain barrier (see, e.g., PCT Publication No. WO 89/10134, published April 25, 1988), hybridization-triggered cleavage agents (see, e.g., Krol et 20 al., 1988, BioTechniques 6:958-976) or intercalating agents (see, e.g., Zon, 1988, Pharm. Res. 5:539-549).

In a preferred aspect of the invention, a lats antisense oligonucleotide is provided, preferably of single-stranded DNA. In a most preferred aspect, such an 25 oligonucleotide comprises a sequence antisense to the sequence encoding an SH3 binding domain or a kinase domain of a lats protein, most preferably, of a human lats protein. The oligonucleotide may be modified at any position on its structure with substituents generally known in the art.

30 The lats antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 35 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, in sine, N6-isopentenyladenine,

1-methylguanine, 1-methyladenine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, 5 beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-10 5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

In another embodiment, the oligonucleotide comprises at least one modified sugar moiety selected from 15 the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a 20 phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the oligonucleotide is an  $\alpha$ -anomeric oligonucleotide. An  $\alpha$ -anomeric oligonucleotide 25 forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641).

The oligonucleotide may be conjugated to another 30 molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

Oligonucleotides of the invention may be synthesized by standard methods known in the art, e.g. by use 35 of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be

synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-5 7451), etc.

In a specific embodiment, the lats antisense oligonucleotide comprises catalytic RNA, or a ribozyme (see, e.g., PCT International Publication WO 90/11364, published October 4, 1990; Sarver et al., 1990, Science 247:1222-1225).

10 In another embodiment, the oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al., 1987, FEBS Lett. 215:327-330).

In an alternative embodiment, the lats antisense  
15 nucleic acid of the invention is produced intracellularly by transcription from an exogenous sequence. For example, a vector can be introduced in vivo such that it is taken up by a cell, within which cell the vector or a portion thereof is transcribed, producing an antisense nucleic acid (RNA) of the  
20 invention. Such a vector would contain a sequence encoding the lats antisense nucleic acid. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology  
25 methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the lats antisense RNA can be by any promoter known in the art to act in mammalian, preferably human, cells.  
30 Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980, Cell 22:787-  
35 797), the herpes thymidine kinase promoter (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445), the

regulatory sequences of the metallothionein gene (Brinster et al., 1982, Nature 296:39-42), etc.

The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a *lats* gene, preferably a human *lats* gene. However, absolute complementarity, although preferred, is not required. A sequence "complementary to at least a portion of an RNA," as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded *lats* antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with a *lats* RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

#### 5.8.2.1.2. THERAPEUTIC USE OF *LATS* ANTISENSE NUCLEIC ACIDS

The *lats* antisense nucleic acids can be used to treat (or prevent) disorders of a cell type that expresses, or preferably overexpresses, *lats*. In a specific embodiment, such a disorder is a growth deficiency. In a preferred embodiment, a single-stranded DNA antisense *lats* oligonucleotide is used.

Cell types which express or overexpress *lats* RNA can be identified by various methods known in the art. Such methods include but are not limited to hybridization with a *lats*-specific nucleic acid (e.g. by Northern hybridization, dot blot hybridization, *in situ* hybridization), observing the ability of RNA from the cell type to be translated *in vitro* into *lats*, immunoassay, etc. In a preferred aspect, primary tissue from a patient can be assayed for *lats* expression

prior to treatment, e.g., by immunocytochemistry or *in situ* hybridization.

Pharmaceutical compositions of the invention (see Section 5.10), comprising an effective amount of a *lats* antisense nucleic acid in a pharmaceutically acceptable carrier, can be administered to a patient having a disease or disorder which is of a type that expresses or overexpresses *lats* RNA or protein.

The amount of *lats* antisense nucleic acid which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. Where possible, it is desirable to determine the antisense cytotoxicity of the tumor type to be treated *in vitro*, and then in useful animal model systems prior to testing and use in humans.

In a specific embodiment, pharmaceutical compositions comprising *lats* antisense nucleic acids are administered via liposomes, microparticles, or microcapsules. In various embodiments of the invention, it may be useful to use such compositions to achieve sustained release of the *lats* antisense nucleic acids. In a specific embodiment, it may be desirable to utilize liposomes targeted via antibodies to specific identifiable tumor antigens (Leonetti et al., 1990, Proc. Natl. Acad. Sci. U.S.A. 87:2448-2451; Renneisen et al., 1990, J. Biol. Chem. 265:16337-16342).

Additional methods that can be adapted for use to deliver a *lats* antisense nucleic acid are described in Section 5.8.1.4.

30

#### 5.9. DEMONSTRATION OF THERAPEUTIC OR PROPHYLACTIC UTILITY

The Therapeutics of the invention are preferably tested *in vitro*, and then *in vivo* for the desired therapeutic or prophylactic activity, prior to use in humans. For example, *In vitro* assays which can be used to determine whether administration of a specific Therapeutic is

35

indicated, include *in vitro* cell culture assays in which a patient tissue sample is grown in culture, and exposed to or otherwise administered a Therapeutic, and the effect of such Therapeutic upon the tissue sample is observed. In one  
5 embodiment, where the patient has a malignancy, a sample of cells from such malignancy is plated out or grown in culture, and the cells are then exposed to a Therapeutic. A Therapeutic which inhibits survival or growth of the malignant cells is selected for therapeutic use *in vivo*.  
10 Many assays standard in the art can be used to assess such survival and/or growth; for example, cell proliferation can be assayed by measuring <sup>3</sup>H-thymidine incorporation, by direct cell count, by detecting changes in transcriptional activity of known genes such as proto-oncogenes (e.g., *fos*, *myc*) or  
15 cell cycle markers; cell viability can be assessed by trypan blue staining, differentiation can be assessed visually based on changes in morphology, etc.

In another embodiment, a Therapeutic is indicated for use which exhibits the desired effect, inhibition or  
20 promotion of cell growth, upon a patient cell sample from tissue having or suspected of having a hyper- or hypoproliferative disorder, respectively. Such hyper- or hypoproliferative disorders include but are not limited to those described in Sections 5.8.1 through 5.8.3 *infra*.

25 In another specific embodiment, a Therapeutic is indicated for use in treating cell injury or a degenerative disorder (see Section 5.8.2) which exhibits *in vitro* promotion of growth/proliferation of cells of the affected patient type. Regarding nervous system disorders, see also  
30 Section 5.8.2.1 for assays that can be used.

In various specific embodiments, *in vitro* assays can be carried out with representative cells of cell types involved in a patient's disorder, to determine if a Therapeutic has a desired effect upon such cell types.

35 In another embodiment, cells of a patient tissue sample suspected of being pre-neoplastic are similarly plated out or grown *in vitro*, and exposed to a Therapeutic. The



Therapeutic which results in a cell phenotype that is more normal (i. ., less representative of a pre-neoplastic state, neoplastic state, malignant state, or transformed phenotype) is selected for therapeutic use. Many assays standard in the art can be used to assess whether a pre-neoplastic state, neoplastic state, or a transformed or malignant phenotype, is present. For example, characteristics associated with a transformed phenotype (a set of *in vitro* characteristics associated with a tumorigenic ability *in vivo*) include a more rounded cell morphology, looser substratum attachment, loss of contact inhibition, loss of anchorage dependence, release of proteases such as plasminogen activator, increased sugar transport, decreased serum requirement, expression of fetal antigens, disappearance of the 250,000 dalton surface protein, etc. (see Luria et al., 1978, *General Virology*, 3d Ed., John Wiley & Sons, New York pp. 436-446).

In other specific embodiments, the *in vitro* assays described *supra* can be carried out using a cell line, rather than a cell sample derived from the specific patient to be treated, in which the cell line is derived from or displays characteristic(s) associated with the malignant, neoplastic or pre-neoplastic disorder desired to be treated or prevented, or is derived from the cell type upon which an effect is desired, according to the present invention.

Compounds for use in therapy can be tested in suitable animal model systems prior to testing in humans, including but not limited to rats, mice, chicken, cows, monkeys, rabbits, etc. For *in vivo* testing, prior to administration to humans, any animal model system known in the art may be used.

#### 5.10. THERAPEUTIC/PROPHYLACTIC ADMINISTRATION AND COMPOSITIONS

The invention provides methods of treatment (and prophylaxis) by administration to a subject of an effective amount of a Therapeutic of the invention. In a preferred aspect, the Therapeutic is substantially purified. The

subject is preferably an animal, including but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, tc., and is preferably a mammal, and most preferably human. In a specific embodiment, a non-human mammal is the subject.

- 5 Formulations and methods of administration that can be employed when the Therapeutic comprises a nucleic acid are described in Sections 5.8.1.4 and 5.8.2.2 above; additional appropriate formulations and routes of administration can be selected from among those described hereinbelow.
- 10 Various delivery systems are known and can be used to administer a Therapeutic of the invention, e.g., encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the Therapeutic, receptor-mediated endocytosis (see, e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), construction of a Therapeutic
- 15 nucleic acid as part of a retroviral or other vector, etc. Methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The
- 20 compounds may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents.
- 25 Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an
- 30 intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.
- 35 In a specific embodiment, it may be desirable to administer the pharmaceutical compositions of the invention locally to the area in need of treatment; this may be

achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, 5 or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. In one embodiment, administration can be by direct injection at the site (or former site) of a malignant tumor or neoplastic or pre- 10 neoplastic tissue.

In another embodiment, the Therapeutic can be delivered in a vesicle, in particular a liposome (see Langer, Science 249:1527-1533 (1990); Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein 15 and Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, *ibid.*, pp. 317-327; see generally *ibid.*)

In yet another embodiment, the Therapeutic can be delivered in a controlled release system. In one embodiment, a pump may be used (see Langer, *supra*; Sefton, CRC Crit. Ref. 20 Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989)). In another embodiment, polymeric materials can be used (see Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974); Controlled 25 Drug Bioavailability, Drug Product Design and Performance, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, J. Macromol. Sci. Rev. Macromol. Chem. 23:61 (1983); see also Levy et al., Science 228:190 (1985); During et al., Ann. Neurol. 25:351 (1989); Howard et al., J. Neurosurg. 30 71:105 (1989)). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, i.e., the brain, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in Medical Applications of Controlled Release, *supra*, vol. 2, pp. 35 115-138 (1984)).

Other controlled release systems are discussed in the review by Langer (Science 249:1527-1533 (1990)).

In a specific embodiment where the Therapeutic is a nucleic acid encoding a protein Therapeutic, the nucleic acid can be administered *in vivo* to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by use of a retroviral vector (see U.S. Patent No. 4,980,286), or by direct injection, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see e.g., Joliot et al., 1991, Proc. Natl. Acad. Sci. USA 88:1864-1868), etc. Alternatively, a nucleic acid Therapeutic can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a Therapeutic, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium

chlorid , dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the Therapeutic, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation should suit the mode of administration.

In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

The Therap utics of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with free amino groups such as those derived from hydrochloric, phosphoric, acetic, oxalic, 5 tartaric acids, etc., and those formed with free carboxyl groups such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

10 The amount of the Therapeutic of the invention which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. In addition, in vitro assays may 15 optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each 20 patient's circumstances. However, suitable dosage ranges for intravenous administration are generally about 20-500 micrograms of active compound per kilogram body weight. Suitable dosage ranges for intranasal administration are generally about 0.01 pg/kg body weight to 1 mg/kg body 25 weight. Effective doses may be extrapolated from dose-response curves derived from in vitro or animal model test systems.

Suppositories generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations 30 preferably contain 10% to 95% active ingredient.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) 35 can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmac uticals or biological products, which notice reflects

approval by the agency of manufacture, use or sale for human administration.

5.11. ADDITIONAL USE OF INHIBITION OF LATS  
FUNCTION TO PROMOTE INCREASED GROWTH

5 Inhibition of lats function (e.g., by administering a compound that inhibits lats function as described in Sections 5.8.2 through 5.8.2.1.2 above), has utility that is not limited to therapeutic or prophylactic applications. For  
10 example, lats function can be inhibited in order to increase growth of animals (e.g., cows, horses, pigs, goats, deer, chickens) and plants (particularly edible plants, e.g., tomatoes, melons, lettuce, carrots, potatoes, and other vegetables), particularly those that are food or material  
15 sources. For example, antisense inhibition (preferably where the lats antisense nucleic acid is under the control of a tissue-specific promoter) can be used in plants or animals to increase growth where desired (e.g., in the fruit or muscle). For example, a lats antisense nucleic acid under the control  
20 of a temperature-sensitive promoter can be administered to a plant or animal, and the desired portion of the (or the entire) plant or animal can be subjected to heat in order to induce antisense nucleic acid production, resulting lats inhibition, and resulting cell proliferation. In other  
25 embodiments, chemical mutagenesis, or homologous recombination with an insertionally inactivated lats gene (see Capecchi, 1989, Science 244:1288-1292 and Section 5.14 *infra*) can be carried out to reduce or destroy endogenous lats function, in order to achieve increased growth.  
30 Suitable methods, modes of administration and compositions, that can be used to inhibit lats function are described in Sections 5.8.2 through 5.8.2.1.2, above. Methods to make plants recombinant are commonly known in the art and can be used. Regarding methods of plant transformation (e.g., for  
35 transformation with a lats antisense nucleic acid or with a sequence encoding a lats derivative that is a dominant-negative kinase), see e.g., Valvekens et al., 1988, Proc.

Natl. Acad. Sci. USA 85:5536-5540. Regarding methods of targeted gene inactivation in plants (e.g., to inactivate lats), see e.g., Miao and Lam, 1995, The Plant J. 7:359-365.

Inhibition of lats function can also have uses *in vitro*, e.g., to expand cells *in vitro*, including but not limited to stem cells, progenitor cells, muscle cells, fibroblasts, liver cells, etc., e.g., to grow cells/tissue *in vitro* prior to administration to a patient (preferably a patient from which the cells were derived), etc.

10

#### 5.12. ADDITIONAL USE OF INHIBITION OF LATS FUNCTION TO INHIBIT CELLULAR SENESENCE

Inhibition of lats function (e.g., by administering a compound that inhibits lats function as described in Sections 5.8.2 through 5.8.2.1.2 above), also has utility in the inhibition of cellular senescence. Thus, inhibition of lats function can be carried out to delay or prevent the onset of cellular senescence, *in vivo* or *in vitro*. In a specific embodiment, cellular senescence is delayed or prevented without incurring the onset of cell malignancy or its *in vitro* correlate, a transformed phenotype.

Thus, for example, a lats antagonist (e.g., anti-lats antibody, lats derivatives or analogs that are dominant-negative kinases; lats antisense nucleic acids, etc.) can be administered to a subject to inhibit or prevent aging or cell death or the effects of aging or cell death (e.g., in the skin, wrinkling, loss of elasticity, less uniform skin tone; in the skin and elsewhere, loss of known characteristics of proper physiological function such as expression of characteristic antigens, secreted molecules, etc.) In one embodiment, a lats antagonist is applied topically, e.g., in a cream or gel, to the skin of the subject. In another embodiment, a lats antagonist is injected, e.g., intradermally, intraperitoneally, or intramuscularly.

In a specific embodiment, a lats antagonist is contacted with cells grown in culture, e.g., by addition of the antagonist to the culture medium or by adsorption of the



antagonist to the culture plate or flask prior to seeding of the cells, in order to inhibit or delay senescence in vitro, e.g., to delay "crisis" phase. For example, such a method can be carried out in order to lengthen the time that cells can be kept alive in vitro, e.g., in order to facilitate conducting studies of the toxicity of a compound (e.g., a lead drug candidate) upon such cells, to study the effect of a molecule upon cell function, and, generally, to study the function of such cells. Such cells include but are not limited to neurons of the central nervous system (e.g., hippocampal, hypothalamic) or peripheral nervous system, glial cells, fibroblasts, kidney cells, liver cells, heart cells, muscle cells, endothelial cells, melanocytes, and hematopoietic cells such as T and B lymphocytes, macrophages, granulocytes, and mast cells.

In vitro assays of senescence are well known in the art and can be used to screen potential lats antagonists prior to use in this aspect of the invention (see, e.g., Hubbard and Ozer, 1995, "Senescence and immortalization of human cells," in Cell Growth and Apoptosis, A Practical Approach, Ch. 12, Studzinski, G.P. (ed.), Oxford University Press, Inc., New York, NY, pp. 229-248.

#### 5.13. DIAGNOSIS AND SCREENING

Lats proteins, analogues, derivatives, and subsequences thereof, lats nucleic acids (and sequences complementary thereto), anti-lats antibodies, have uses in diagnostics. Such molecules can be used in assays, such as immunoassays, to detect, prognose, diagnose, or monitor various conditions, diseases, and disorders affecting lats expression, or monitor the treatment thereof. In particular, such an immunoassay is carried out by a method comprising contacting a sample derived from a patient with an anti-lats antibody under conditions such that immunospecific binding can occur, and detecting or measuring the amount of any immunospecific binding by the antibody. In a specific aspect, such binding of antibody, in tissue sections, can be

used to detect aberrant lats localization or aberrant (e.g., low or absent) levels of lats. In a specific embodiment, antibody to lats can be used to assay in a patient tissue or serum sample for the presence of lats where an aberrant level of lats is an indication of a diseased condition. By "aberrant levels," is meant increased or decreased levels relative to that present, or a standard level representing that present, in an analogous sample from a portion of the body or from a subject not having the disorder.

10 The immunoassays which can be used include but are not limited to competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin  
15 reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays, protein A immunoassays, to name but a few.

Lats genes and related nucleic acid sequences and  
20 subsequences, including complementary sequences, can also be used in hybridization assays. Lats nucleic acid sequences, or subsequences thereof comprising about at least 8 nucleotides, can be used as hybridization probes. Hybridization assays can be used to detect, prognose,  
25 diagnose, or monitor conditions, disorders, or disease states associated with aberrant changes in lats expression and/or activity as described supra. In particular, such a hybridization assay is carried out by a method comprising contacting a sample containing nucleic acid with a nucleic  
30 acid probe capable of hybridizing to lats DNA or RNA, under conditions such that hybridization can occur, and detecting or measuring any resulting hybridization.

In specific embodiments, diseases and disorders involving overproliferation of cells can be diagnosed, or  
35 their suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by detecting decreased levels of lats protein, lats RNA, or lats

functional activity (e.g., kinase activity, SH3 domain-binding activity, tc.), or by detecting mutations in lats RNA, DNA or protein (e.g., translocations in lats nucleic acids, truncations in the lats gene or protein, changes in nucleotide or amino acid sequence relative to wild-type lats) that cause decreased expression or activity of lats. Such diseases and disorders include but are not limited to those described in Section 5.8.1 and its subsections. By way of example, levels of lats protein can be detected by immunoassay, levels of lats RNA can be detected by hybridization assays (e.g., Northern blots, dot blots), lats kinase activity can be measured by kinase assays commonly known in the art, lats binding to an SH3 domain-containing protein can be done by binding assays commonly known in the art, translocations and point mutations in lats nucleic acids can be detected by Southern blotting, RFLP analysis, PCR using primers that preferably generate a fragment spanning at least most of the lats gene, sequencing of the lats genomic DNA or cDNA obtained from the patient, etc.

In a preferred embodiment, levels of lats mRNA or protein in a patient sample are detected or measured, in which decreased levels indicate that the subject has, or has a predisposition to developing, a malignancy or hyperproliferative disorder; in which the decreased levels are relative to the levels present in an analogous sample from a portion of the body or from a subject not having the malignancy or hyperproliferative disorder, as the case may be.

In another specific embodiment, diseases and disorders involving a deficiency in cell proliferation or in which cell proliferation is desirable for treatment, are diagnosed, or their suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by detecting increased levels of lats protein, lats RNA, or lats functional activity (e.g., kinase activity, SH3 domain binding activity, etc.), or by detecting mutations in lats RNA, DNA or protein (e.g., translocations in lats

nucl ic acids, truncations in the gene or protein, changes in nucl otide or amino acid sequence relative to wild-type lats) that cause incr ased expr ssion or activity of lats. Such diseases and disorders include but are not limited to those 5 described in Section 5.8.2 and its subsections. By way of example, levels of lats protein, levels of lats RNA, lats kinase activity, lats binding activity, and the presence of translocations or point mutations can be determined as described above.

10 In a specific embodiment, levels of lats mRNA or protein in a patient sample are detected or measured, in which increased levels indicate that the subject has, or has a predisposition to developing, a growth deficiency or degenerative or hypoproliferative disorder; in which the 15 increased levels are relative to the levels present in an analogous sample from a portion of the body or from a subject not having the growth deficiency, degenerative, or hypoproliferative disorder, as the case may be.

Kits for diagnostic use are also provided, that 20 comprise in one or more containers an anti-lats antibody, and, optionally, a labeled binding partner to the antibody. Alternatively, the anti-lats antibody can be labeled (with a detectable marker, e.g., a chemiluminescent, enzymatic, fluorescent, or radioactive moiety). A kit is also provided 25 that comprises in one or more containers a nucleic acid probe capable of hybridizing to lats RNA. In a specific embodiment, a kit can comprise in one or more containers a pair of primers (e.g., each in the size range of 6-30 nucleotides) that are capable of priming amplification [e.g., 30 by polymerase chain reaction (see e.g., Innis et al., 1990, PCR Protocols, Academic Press, Inc., San Diego, CA), ligase chain reaction (see EP 320,308) use of Q $\beta$  replicase, cyclic probe reaction, or other methods known in the art] under appropriate reaction conditions of at least a portion of a 35 lats nucleic acid. A kit can optionally further comprise in a container a predetermined amount of a purified lats protein or nucleic acid, e.g., for use as a standard or control.

#### 5.14. SCREENING FOR LATS AGONISTS AND ANTAGONISTS

Lats nucleic acids, proteins, and derivatives also have uses in screening assays to detect molecules that specifically bind to lats nucleic acids, proteins, or derivatives and thus have potential use as agonists or antagonists of lats, in particular, molecules that thus affect cell proliferation. In a preferred embodiment, such assays are performed to screen for molecules with potential utility as anti-cancer drugs or lead compounds for drug development. The invention thus provides assays to detect molecules that specifically bind to lats nucleic acids, proteins, or derivatives. For example, recombinant cells expressing lats nucleic acids can be used to recombinantly produce lats proteins in these assays, to screen for molecules that bind to a lats protein. Molecules (e.g., putative binding partners of lats) are contacted with the lats protein (or fragment thereof) under conditions conducive to binding, and then molecules that specifically bind to the lats protein are identified. Similar methods can be used to screen for molecules that bind to lats derivatives or nucleic acids. Methods that can be used to carry out the foregoing are commonly known in the art.

By way of example, diversity libraries, such as random or combinatorial peptide or nonpeptide libraries can be screened for molecules that specifically bind to lats. Many libraries are known in the art that can be used, e.g., chemically synthesized libraries, recombinant (e.g., phage display libraries), and in vitro translation-based libraries.

Examples of chemically synthesized libraries are described in Fodor et al., 1991, Science 251:767-773; Houghten et al., 1991, Nature 354:84-86; Lam et al., 1991, Nature 354:82-84; Medynski, 1994, Bio/Technology 12:709-710; Gallop et al., 1994, J. Medicinal Chemistry 37(9):1233-1251; Ohlmeyer et al., 1993, Proc. Natl. Acad. Sci. USA 90:10922-10926; Erb et al., 1994, Proc. Natl. Acad. Sci. USA 91:11422-11426; Houghten et al., 1992, Biotechniques 13:412; Jayawickreme et al., 1994, Proc. Natl. Acad. Sci. USA

91:1614-1618; Salmon et al., 1993, Proc. Natl. Acad. Sci. USA 90:11708-11712; PCT Publication No. WO 93/20242; and Brenner and Lerner, 1992, Proc. Natl. Acad. Sci. USA 89:5381-5383.

Examples of phag display libraries are described in Scott and Smith, 1990, Science 249:386-390; Devlin et al., 1990, Science, 249:404-406; Christian, R.B., et al., 1992, J. Mol. Biol. 227:711-718; ~~Lenstra, 1992, J. Immunol. Meth. 152:149-157; Kay et al., 1993, Gene 128:59-65; and PCT~~ Publication No. WO 94/18318 dated August 18, 1994.

10 In vitro translation-based libraries include but are not limited to those described in PCT Publication No. WO 91/05058 dated April 18, 1991; and Mattheakis et al., 1994, Proc. Natl. Acad. Sci. USA 91:9022-9026.

By way of examples of nonpeptide libraries, a benzodiazepine library (see e.g., Bunin et al., 1994, Proc. Natl. Acad. Sci. USA 91:4708-4712) can be adapted for use. Peptoid libraries (Simon et al., 1992, Proc. Natl. Acad. Sci. USA 89:9367-9371) can also be used. Another example of a library that can be used, in which the amide functionalities in peptides have been permethylated to generate a chemically transformed combinatorial library, is described by Ostresh et al. (1994, Proc. Natl. Acad. Sci. USA 91:11138-11142).

Screening the libraries can be accomplished by any of a variety of commonly known methods. See, e.g., the following references, which disclose screening of peptide libraries: Parmley and Smith, 1989, Adv. Exp. Med. Biol. 251:215-218; Scott and Smith, 1990, Science 249:386-390; Fowlkes et al., 1992, BioTechniques 13:422-427; Oldenburg et al., 1992, Proc. Natl. Acad. Sci. USA 89:5393-5397; Yu et al., 1994, Cell 76:933-945; Staudt et al., 1988, Science 241:577-580; Bock et al., 1992, Nature 355:564-566; Tuerk et al., 1992, Proc. Natl. Acad. Sci. USA 89:6988-6992; Ellington et al., 1992, Nature 355:850-852; U.S. Patent No. 5,096,815, U.S. Patent No. 5,223,409, and U.S. Patent No. 5,198,346, all to Ladner et al.; Rebar and Pabo, 1993, Science 263:671-673; and PCT Publication No. WO 94/18318.

In a specific embodiment, screening can be carried out by contacting the library members with a lats protein (or nucleic acid or derivativ ) immobilized on a solid phase and harv sting those library members that bind to the protein (or 5 nucleic acid or derivative). Examples of such screening methods, termed "panning" techniques are described by way of ~~example in Parmley and Smith, 1988, Gene 73:305-318; Fowlkes et al., 1992, BioTechniques 13:422-427; PCT Publication No. WO 94/18318; and in references cited hereinabove.~~

10 In another embodiment, the two-hybrid system for selecting interacting proteins in yeast (Fields and Song, 1989, Nature 340:245-246; Chien et al., 1991, Proc. Natl. Acad. Sci. USA 88:9578-9582) can be used to identify molecules that specifically bind to a lats protein or 15 derivative.

In addition, *Drosophila* can be used as a model system in order to detect genes that phenotypically interact with lats. For example, overexpression of lats in *Drosophila* eye leads to a smaller and rougher eye. Mutagenesis of the 20 fly genome can be performed, followed by selecting flies in which the mutagenesis has resulted in suppression or enhancement of the small rough eye phenotype; the mutated genes in such flies are likely to encode proteins that interact/bind with lats.

25

#### 5.15. ANIMAL MODELS

The invention also provides animal models.

In one embodiment, animal models for diseases and disorders involving cell overproliferation (e.g., as 30 described in Section 5.8.1) are provided. Such an animal can be initially produced by promoting homologous recombination between a lats gene in its chromosome and an exogenous lats gene that has been rendered biologically inactive (preferably by insertion of a heterologous sequence, e.g., an antibiotic 35 resistance gene). In a preferred aspect, this homologous recombination is carried out by transforming embryo-derived stem (ES) cells with a vector containing the insertionally

inactivated *lats* gene, such that homologous recombination occurs, followed by injecting the ES cells into a blastocyst, and implanting the blastocyst into a foster mother, followed by the birth of the chimeric animal ("knockout animal") in which a *lats* gene has been inactivated (see Capecchi, 1989, Science 244:1288-1292). The chimeric animal can be bred to produce additional knockout animals. Such animals can be mice, hamsters, sheep, pigs, cattle, etc., and are preferably non-human mammals. In a specific embodiment, a knockout mouse is produced.

Such knockout animals are expected to develop or be predisposed to developing diseases or disorders involving cell overproliferation (e.g., malignancy) and thus can have use as animal models of such diseases and disorders, e.g., to screen for or test molecules (e.g., potential anti-cancer therapeutics) for the ability to inhibit overproliferation (e.g., tumor formation) and thus treat or prevent such diseases or disorders.

In a different embodiment of the invention, transgenic animals that have incorporated and express a functional *lats* gene have use as animal models of diseases and disorders involving deficiencies in cell proliferation or in which cell proliferation is desired. Such animals can be used to screen for or test molecules for the ability to promote proliferation and thus treat or prevent such diseases and disorders.

#### 5.16. METHODS OF IDENTIFYING TUMOR SUPPRESSOR GENES AND OTHER GENES WITH IDENTIFIABLE PHENOTYPES

The invention also provides methods of identifying a tumor suppressor gene (or potential tumor suppressor gene) comprising identifying an overproliferation phenotype in a genetic mosaic, and isolating a gene that is mutated in cells exhibiting the overproliferation phenotype. The genetic mosaic is achieved by induction of somatic cells in an animal that is heterozygous for an induced mutation to become



homozygous for the mutation, at any desired developmental stage. The mutation can be induced by any known method, e.g., X-ray exposure or chemical mutation exposure or insertion of a transposable element (e.g., P-element). A genetic mosaic is produced by induction of homozygosity by mitotic recombination between homologous arms of both parental chromosomes, which is achieved using a site-specific recombination system [a sequence capable of expressing a site-specific recombinase; and its target sites (sequences at which the recombinase promotes recombination)], that have been inserted in the homozygous arms of both parental chromosomes. The target sites are preferably inserted close to the centromere on each chromosome arm (the closer to the centromere, the more preferred), so that mitotic recombination events will result in cells being homozygous for the mutation located on the chromosome arm distal to the insertion of the target site. For example, an FLP recombinase can be used with FRT target sites; Cre recombinase can be used with lox target sites. The recombinase coding sequence, used to express recombinase, preferably, but need not be, intrachromosomally situated. For at least one chromosome, the target sites are intrachromosomally inserted on the homologous arms of both parental (maternal and paternal) chromosomes.

The genetic mosaic can be an animal, e.g., mouse, hamster, sheep, pig, cow, *Drosophila*, etc., and is preferably a non-human mammal.

In a specific embodiment relating to the production of a non-human mammal that is a genetic mosaic, a recombinase target site is introduced onto one arm of a chromosome in an embryo-derived stem cell (ES). The target site can be introduced into the cell by homologous recombination (by use of flanking sequences from the desired site of intrachromosomal integration) or by random integration resulting from cell transformation (e.g., by transfection, electroporation), etc. This ES is then injected into a blastocyst, the blastocyst is implanted into a foster mother,

followed by birth of the recombinant animal. This mammal is bred to a wild-type female, to produce siblings. Siblings carrying the target site insertion are mated, and offspring carrying the target site on the homologous arms of both parental chromosomes are isolated ("the target strain"). A target strain member is then mutagenized and mated with a non-mutagenized target strain member of the opposite sex (preferably also carrying a recombinant nucleic acid encoding and capable of expressing a recombinase that promotes recombination at the target sites), to obtain a target strain member that is heterozygous for the mutation. Provision of the recombinase (by expression) in mitotically active cells of a developing animal or an adult animal promotes mitotic recombination between the homologous arms of the parental chromosomes, resulting in a cell that is homozygous for the mutation. Cells that display a mutant phenotype by virtue of their being homozygous for the mutation are then detected, and the mutant gene can be genetically mapped by any known method, and can be isolated.

20 In a *Drosophila* animal, a site-specific recombination system can be introduced by use of P-element-mediated insertions.

In one embodiment, target sites are introduced onto homologous arms of both of a set of parental chromosomes, for one chromosome. In another embodiment, target sites are introduced onto homologous arms of both of a set of parental chromosomes, for a plurality of chromosomes.

The recombinase can be under the control of a constitutive (e.g., phosphorylated kinase promoter) or inducible (e.g., heat shock promoter) or tissue-specific promoter. The recombinase can be expressed episomally (e.g., from a plasmid) or chromosomally. Once the recombination system is introduced into the animal, genetic mosaicism is produced by the activity of the recombinase (which promotes recombination at the target sites).

In a specific embodiment, an animal is used that contains a recombinant nucleic acid encoding an FLP

recombinase (Broach and Hicks, 1980, Cell 21:501-508) such that it is expressible by a cell of the animal, and intrachromosomal insertions of an FRT site on the homologous arms of both parental chromosomes; and genetic mosaicism is produced by inducing mitotic recombination between the FRT sites on the homologous chromosome arms after FLP recombinase expression (e.g., by heat shock, when expression of the FLP recombinase is under the control of a heat shock promoter).

In another specific embodiment, an animal is used that contains a recombinant nucleic acid encoding a Cre recombinase (Sauer and Henderson, 1988, Proc. Natl. Acad. Sci. USA 85:5166-5170) such that it is expressible by a cell of the animal, and intrachromosomal insertions of a lox site on homologous arms of both parental chromosomes; and genetic mosaicism is produced by inducing mitotic recombination between the lox sites on the homologous chromosome arms after Cre recombinase expression.

The animal may optionally further comprise intrachromosomal insertions of marker genes (comprising a sequence encoding a protein containing a reporter group such as an epitope tag), to facilitate confirmation and/or monitoring of recombination events. For example, in a non-human mammal, a marker gene (e.g., lacZ) operably linked to a constitutive promoter can be inserted, on the same chromosome arm as that carrying the target site and the induced mutation.

In a specific embodiment, the overproliferation phenotype is the formation of overproliferated outgrowth tissue in a non-position-dependent fashion. In another specific embodiment, the overproliferation phenotype is the formation of a normal structure of larger than normal size.

The above-described genetic mosaics have uses not only in identifying tumor suppressor genes, but, more generally, in identifying genes with an identifiable phenotype, i.e., those genes which in mutated form cause an observable mutant phenotype to be displayed in the genetic mosaic.

In another embodiment, the invention provides a method of identifying genes with an observable mutant phenotype by use of human (or other animal) tissue culture cells that have incorporated a site-specific recombination system such as described above. The site-specific recombination system can be introduced by methods such as described above, so as to introduce a recombinant source of recombinase and effect intrachromosomal insertions of the recombinase target sites on the homologous arms of both of a set of parental chromosomes, for one or more chromosomes. In a preferred aspect relating to this use of culture cells, the recombinase target sites are ligated to a selectable marker (e.g., an antibiotic resistance gene), and cells are obtained that have the target sites on each of the homologous chromosome arms, by selecting under selection conditions of relatively high stringency (e.g., by increasing the antibiotic concentration in the cell medium). As with the use of genetic mosaics as described above, once mitotic recombination is induced between the target sites on the homologous chromosome arms, one then identifies cells displaying a mutant phenotype, and recovers a gene mutated in cells exhibiting the mutant phenotype. For example, a potential tumor suppressor gene can be identified by isolating a gene that is mutated in cultured cells exhibiting a transformed phenotype.

6. IDENTIFYING TUMOR SUPPRESSORS IN  
GENETIC MOSAICS: THE *DROSOPHILA* LATS  
GENE ENCODES A PUTATIVE PROTEIN KINASE

We have identified recessive overproliferation mutations by screening and examining clones of mutant cells in genetic mosaics of the fruitfly *Drosophila melanogaster* (Fig. 1A). Flies that carry small groups of somatic cells mutated for negative regulators of cell proliferation or tumor suppressors are viable, yet the overproliferated mutant tissues can be readily identifiable.

One way to generate mosaic animals is to induce mitotic recombination in developing heterozygous individuals (Fig. 1B). Recently, it was found that the site-specific recombination system from yeast, the FLP recombinase and its target site FRT, can be used to induce high frequency of mitotic recombination in *Drosophila* (Golic and Lindquist, 1989, Cell 59:499-509; Golic, 1991, Science 252:958-961). To produce and analyze genetic mosaics, a series of special *Drosophila* strains were constructed, containing the FLP/FRT recombination system on genetically marked chromosomes (Xu and Rubin, 1993, Development 117:1223-1237). Using these strains, high frequencies of mosaicism can be produced for more than 95% of the *Drosophila* genes. We have used these strains to identify overproliferation mutations in mosaic animals.

Our results show that screening for overproliferation mutations in mosaic animals is a powerful way to identify negative regulators of cell proliferation and potential tumor suppressor genes. One of the identified genes, large tumor suppressor (*lats*), has been cloned, and encodes a predicted novel protein kinase. Mutations in *lats* cause dramatic overproliferation phenotypes and various developmental defects in both mosaic animals and homozygous mutants.

25

#### 6.1. MATERIALS AND METHODS

##### Genetics

Fly stocks and crosses were grown on standard medium at 25°C unless otherwise indicated. The F1 mosaic screens were modified from the one described in Xu and Rubin (1993, Development 117:1223-1237) and in Xu and Harrison (1994, Methods in Cell Biology 44:655-682). Briefly, the F1 mosaic individuals were produced from three crosses: Mutagenized *y w hsFLP1; P[ry<sup>+</sup>; hs-neo; FRT]40A* males were mated to the *y w hsFLP1; P[ry<sup>+</sup>; y<sup>+</sup>]25F, P[mini-w<sup>+</sup>; hs-NM]31E, P[ry<sup>+</sup>; hs-neo; FRT]40A females. Mutagenized *y w hsFLP1; P[ry<sup>+</sup>; hs-neo; FRT]42D* males were mated to the *y w hsFLP1;**

35

*P[ry<sup>+</sup>; hs-neo; FRT]42D, P[ry<sup>+</sup>; y<sup>+</sup>]44B, P[mini-w<sup>+</sup>; hs-NM]46F/CyO* f mal s. Finally, mutagenized *y w hsFLP1; P[ry<sup>+</sup>; hs-neo; FRT]82B* males were mated to the *y w hsFLP1; P[ry<sup>+</sup>; hs-neo; FRT]82B, P[mini-w<sup>+</sup>; hs- $\pi$ M]87E, Sb<sup>63b</sup>, P[ry<sup>+</sup>; y<sup>+</sup>]96E* 5 females. The male parents were irradiated with X-rays (4000 r) and were removed from the crosses after four days of mating. The eggs from the crosses were collected for every 12 hours and aged for another 30 hours before being incubated in a 38°C water bath for 60 minutes. The F<sub>1</sub> animals were then 10 returned to normal culture conditions until eclosion. About 25,000 F<sub>1</sub> adults from these crosses were examined. Each P-induced lethal mutation was recombined onto one of the FRT-carrying arms using the *neo<sup>R</sup>* and *w* double selection as described in Xu and Harrison (1994, Methods in Cell Biology 15 44:655-682) before examining its clonal phenotype.

The *lats<sup>l</sup>* mutation was meiotically mapped to the right of *claret*. It was further localized to the 100A1-5 region since it complemented *Df(3R)t11<sup>l</sup>(100A2-5; 100C2-3)* and failed to complement *Df(3R)t11<sup>px</sup>(100A1-2; 100B4-5)* and 20 *Df(3R)t11<sup>20</sup>(100A1-3; 100B1-2)*. A saturation genetic screen had previously been performed for this interval, and three lethal complementation groups, *l(3)100Aa*, *l(3)100Ab* and the *zfh-1*, were isolated (Lai et al., 1993, Proc. Natl. Acad. Sci. USA 90:4122-4126). The *lats<sup>l</sup>* mutation failed to 25 complement the EMS-induced mutations in *l(3)100Aa (lats<sup>al-15</sup>)*, but complement mutations in *l(3)100Ab* and *zfh-1*. The clonal phenotypes were examined for *lats<sup>l, Pl, al, a2, a6 and al0</sup>* induced either with the FLP/FRT-marker system or X-ray irradiation.

The *lats<sup>Pl</sup>* allele was recovered from a mosaic male 30 produced from the cross of *y w hsFLP1; P[ry<sup>+</sup>; hs-neo; FRT]82B x y w P[lacZ; w<sup>+</sup>]5; P[ry<sup>+</sup>; hs-neo; FRT]82B/delta2-3, Sb*. The mutant chromosome was cleaned up before performing complementation tests and an excision screen (Robertson et al., 1988, Genetics 118:461-470). Two hundred and fifteen 35 excision lines were established that had lost the *w<sup>+</sup>* gene in the *P[lacZ; w<sup>+</sup>]* element (Bier et al., 1989, Genes Dev.

3:1273-1287). In about 50% of these lines, the pupal lethality had been reverted completely to wild type, indicating the mutant phenotype is caused by the P-element insertion. Five lines were found to cause lethality at late embryonic and/or early first instar larval stages. The remaining lines were found to cause lethality at larval and pupal stages or to produce viable mutant animals. All of these mutant excision lines (except one which is located outside the 100A1-5 region) failed to complement *lats<sup>sl</sup>* and *lats<sup>Pl</sup>*, but do complement mutations in the *zfh-1* and *l(3)100Ab* loci.

The insert in *lats* cDNA A2 was cloned into the pCaSpeR-hs vector (Thummel and Pirrotta, 1992, *Drosophila Inform. Service* 71:150) for germ line transformation. Three of the transformed lines were tested and were able to rescue the lethality of the *lats<sup>sl</sup>/lats<sup>sl</sup>*, *lats<sup>Pl</sup>* and *lats<sup>sl261</sup>* animals after one hour heat shock for every 24 hours during larval and pupal development.

## 20 Histology

Fixation and sectioning (2 mm) of adult *Drosophila* tissues were performed as described (Tomlinson and Ready, 1987, *Dev. Biol.* 123:264-275). Scanning electron microscopy was performed as described (Xu and Artavanis-Tsakonas, 1990, *Genetics* 126:665-677).

## Nucleic Acid Manipulation

A P1 genomic clone (DS02640) mapped in the 100A1-7 region was obtained from the Berkeley *Drosophila* Genome Center (personal communication; Hartl et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:6824-6829). DNA fragments from this P1 clone and genomic DNA obtained by plasmid rescue from the *lats<sup>Pl</sup>* mutant (Bier et al., 1989, *Genes Dev.* 3:1273-1287) were used to isolate several overlapping cosmids including CLT-52 from the genomic library prepared by J. Tamkun. Genomic DNA from +7.5 (*Bgl*III) to -4.2 (*Eco*RI; Fig. 3) was used to screen a total imaginal disc cDNA library prepared by A. Cowman.

Screening approximately 2 million phag yielded three groups of cDNAs (five *lats* cDNAs; fifteen T1 cDNAs; fourteen T2 cDNAs). The sizes of the inserts in the *lats* cDNAs are as follows: 5.6 kb in A2; 5.1 kb in B1; 1.1 kb in 9 and 4; and 0.9 kb in B3.

Genomic DNA from *lats<sup>al</sup>/TM6B*, *lats<sup>al-15</sup>/TM6B*, *lats<sup>pl</sup>/TM6B*, *lats<sup>7-2</sup>/TM6B*, *lats<sup>78</sup>/TM6B*, *lats<sup>el00</sup>/TM6B*, *lats<sup>el19</sup>/TM6B* and *lats<sup>el48</sup>/TM6B* flies was digested with a combination of the *EcoRI*, *BamHI*, *BglII* and *XhoI* restriction enzymes for Southern analysis.

### DNA Sequencing

DNA sequence was determined by the dideoxy chain termination method (Sanger et al., 1977, Proc. Natl. Acad. Sci. USA 74:5463-5467) using Tag polymerase (Perkin Elmer) and Sequenase (U.S. Biochemical Corp.). The sequences of *lats* cDNAs were determined from both strands using templates generated from plasmids containing *EcoRI* fragments inserted into the pBlueScriptII vector. Templates generated from DNase 1 deletion subclones were also used. The complete sequences of cDNAs A2 and 9 were determined; partial sequences were determined for cDNAs B1 and 4. Templates of genomic DNA were generated from plasmids containing *EcoRI* fragments and were sequenced on one strand using synthetic oligonucleotide primers. Mutant DNA from the *lats<sup>al</sup>* allele was amplified with PCR reactions using synthetic oligonucleotide primers and cloned in the pBlueScript II vector for sequencing.

## 6.2. RESULTS

### Screening for Overproliferation Mutations in Mosaic Animals

We have screened individuals carrying clones of cells that were homozygous for either X-ray or P-element induced mutations for overproliferation phenotypes. (Fig. 1B; Materials and Methods). Two types of overproliferation phenotypes were sought: a) Clones of mutant cells form d



ov rproliferat d, outgrowth tissues in a non-position-dependent fashion; b) Clones of mutant cells formed normal structures, but proliferated faster than wild-type cells such that the sizes of the mutant clones were larger than their wt twin-spot clones. Three independent mutations were identified that caused the first type of phenotype (Fig. 2A-2E). A mutation which was allelic to one of the original mutations was later found to cause the second type of phenotype (see below). All three mutations in the first class caused embryonic and/or early larval lethality and they represented single alleles of different loci since they had different chromosome locations. One of them was identified among 215 randomly chosen lethal mutations in which each were caused by a P-element insertion in a different essential gene (Karpen and Spradling, 1992, Genetics 132:737-753; Berkeley *Drosophila* Genome Center, personal communication). In addition to these overproliferation mutations, one P-induced mutation was found to cause both unpatterned outgrowth and duplications of patterned structures in mosaic animals, suggesting that this mutation may not directly affect cell proliferation.

The *lats* Locus Is Defined by a Single  
Complementation Group of Mutations  
That Cause Defects Throughout Development

The mutations caused different levels of overproliferation. One mutation (*lats<sup>1</sup>*) produced much more dramatic overproliferated clones than the ones produced by the other mutations (Fig. 2A, 2B). The *lats* mutant clones induced in first instar larvae can be as large as 1/5 of the body size. Tumorous outgrowth caused by *lats<sup>1</sup>* was found in all the tissues that had been examined including eyes, legs, wings, heads, notums, antenna, and abdominal cuticles. The *lats<sup>1</sup>* mutation was genetically mapped in the 100A1-5 region and the locus was further defined by a single complementation group of over fifty alleles including mutations induced by

X-ray, EMS, P-element insertion and imprecise excision of the P-element (Table 2; Materials and Methods).

TABLE 2

5

The alleles of the *lats* locus\*

Alleles	Phenotypes of homozygous animals	Phenotypes of mutant clones	Representative alleles	No. of alleles
10	Strong	Late embryonic and early 1st instar larval lethal	<i>lats<sup>1</sup></i> , <i>lats<sup>2</sup></i> , <i>lats<sup>3</sup></i>	14
	Medium	Late larval and pupal lethal, normal size of animals	<i>lats<sup>4</sup></i> , <i>lats<sup>124</sup></i>	16
15		Pupal lethal, giant animals	<i>lats<sup>26-1</sup></i>	3
	Weak	Semi-viable and viable: rough eye outgrowth on head, wing held-out, sterile	<i>lats<sup>10</sup></i> , <i>lats<sup>53-2</sup></i>	17

\* The various alleles of the *lats* gene are classified into three main groups as indicated in the left column. Their phenotypes, displayed in either homozygous mutant animals or clones of mutant cells in mosaic animals, are listed in the next two columns respectively. For a given viable or semi-viable allele, the homozygous mutant animals display one, two, three, or all four of the listed phenotypes. Representative alleles and the numbers of alleles for each group are given in the two right columns. The origins of these alleles are described in the Material and Methods.

25 Removing the P-element insertion reverted the lethal chromosome into wild type, indicating the P-element insertion is responsible for the mutant phenotype. Furthermore, five of the imprecise excision lines caused late embryonic and early larval lethality which were stronger than the pupal lethality phenotype caused by the *lats<sup>1</sup>* mutation. These five excision lines failed to complement *lats<sup>1</sup>*, but complemented the mutations in two other complementation groups (*1(3)100Ab* and *zfh-1*) in the 100A1-5 region, indicating that these two genes were not affected by the excision alleles.

35 The *lats* alleles can be classified into three main groups (Table 2). Strong alleles caused homozygous animals to die at a late embryonic stage or shortly after hatching

with no obvious cuticular defect. Mutations in the group of medium alleles cause lethality at different times in larval and pupal development. This group was further divided into two subgroups because three of the excision alleles not only caused pupal lethality, but the sizes of the homozygous mutant animals were also significantly larger than wt animals (Fig. 2C). The weak mutations caused either one or a combination of the following phenotypes: held out wings with broadened blades, rough eye with ventral outgrowth, outgrowth on the dorsal-anterior region of the head and partial to complete sterility (Table 2).

Proliferation defects were observed in both mutant clones in mosaic animals and homozygous mutants. Clones of cells on the head that were homozygous for strong or medium alleles formed unpatterned, overproliferated tissues with many lobes or folds. The mutant cells seemed to be "budding out" of the surface to form new proliferation centers or lobes (Fig. 2A, 2F, 2H). The sizes and the shapes of these mutant cells were very irregular. Cells several times larger than their neighbors were often seen in mutant clones, indicating problematic cell division (Fig. 2F, 2G).

Furthermore, *lats* mutant clones behaved differently from clones mutant for the previously identified *Drosophila* tumor suppressor genes such as *dlg*, *lgl* and *hyd*. The *dlg*, *lgl* or *hyd* mutant cells proliferated slower than wt cells and thus, the mutant clones induced in first instar larvae were competed away during growth and did not form detectable clones in the adults (Bryant, 1987, Experimental and genetic analysis of growth and cell proliferation in *Drosophila* imaginal discs, in "Genetic Regulation of Development," A.R. Liss, New York, pp. 339-372; Woods and Bryant, 1989; Dev. Biol. 134:222-235; Mansfield et al., 1994, Dev. Biol. 165:507-526; Allen Shearn, personal communication). In contrast, the *lats* mutant clones induced at similar developmental stages formed dramatic overproliferated tissues, suggesting the mutant cells proliferated faster than wt cells. Consistent with this notion, clones of cells

mutant for a weak *lats* allele (*lats<sup>al0</sup>*) produced normal looking tissues, but the mutant clones were significantly larger than their wt twin-spot clones. In homozygous animals, the imaginal discs and the central nervous system in many of the pupal lethal mutants were dramatically overproliferated (Fig. 2D, 2E). The discs lost the single layer of epithelial structure and formed multi-layer, deformed tissues. The *lats* overproliferation phenotype was not caused by prevention of differentiation. Cells in the overproliferated mutant clones on the body differentiated and produced bristles and hairs, although the morphologies of these structures were not wild type (Fig. 2I-2L). Careful examination of multiple mutant clones confirmed that *lats* caused mutant cells (w cells in the eye, y bristles and enlarged-base hairs on the body) to overproliferate and did not affect the surrounding wt tissues. Finally, the frequency of overproliferated clones was similar to wt clonal frequency induced with the same FRT element, indicating that loss of the *lats* function alone is sufficient to initiate the overproliferation process.

20

#### Cloning of the *lats* Gene

Genomic DNA from the 100A1-5 region was isolated using probes mapped to this region (Materials and Methods). A restriction map of the relevant genomic region is illustrated in Figure 3. Genomic DNA flanking the P-insertion site (+7.5 to -4.2) was used to screen a total imaginal disc cDNA library. A group of cDNAs corresponding to a 5.7 kb transcript (*lats*) was found to contain sequence from the region where the P-element was inserted (Fig. 3). Two other groups of cDNAs were also isolated (T1 and T2). The 5.7 kb transcript was located in an intron of the T1 gene (Fig. 3). The intron-exon structure of the 5.7 transcription unit was determined by Southern and sequence analysis of the cDNA clones and the corresponding genomic DNA (Materials and Methods). The *zfh-1* gene was found to be located at the left side of the 5.7 kb transcription unit (Fig. 3; Fortini et al., 1991, Mechanisms of Development 34:113-122).

In addition to *lats<sup>PI</sup>*, genomic DNA from the five strong excision alleles was analyzed. All of them deleted exon sequences from the 5.7 kb transcript and, in addition, three of them also deleted sequences in the next transcript 5 (T2; Fig. 3). Furthermore, DNA from the X-ray and EMS induced mutants was analyzed with cDNA probes made from the 5.7 kb, T2 and T1 transcripts. In two cases alterations were detected in the 5.7 kb transcription unit: a 0.4 kb and a 0.3 kb deletions associated with *lats<sup>ai</sup>* and *lats<sup>ad</sup>*, respectively 10 (Fig. 3). The 446 bp deletion in *lats<sup>ai</sup>* was revealed by sequencing. It removed codons 92 to 238 of the open reading frame and caused a frame shift from codon 239 (Fig. 5). Finally, transformants containing a cDNA corresponding to the 5.7 transcript driving by the hsp70 promoter rescued the 15 lethality of both strong and medium *lats* alleles. These findings indicate that the 5.7 kb transcription unit which correspond to the *lats* gene and strong *lats* alleles including *lats<sup>ai</sup>* were either amorphic or nearly amorphic alleles.

## 20 The *lats* Gene Encodes a Putative Protein-Serine/Threonine Kinase

The 5.7 kb *lats* transcript was detected throughout development (Fig. 4) and in both adult males and females (data not shown). In addition, probes from the 5.7 kb 25 transcript also detected a second transcript, which is about 1 kb shorter (4.7 kb), in young embryos (0-4 hrs; Fig. 4) and in adult males and females. Northern analysis showed there was more maternally deposited 4.7 kb transcripts than 5.7 kb transcripts in young embryos (0-2 hrs; Fig. 4). The 5.7 kb 30 transcript became the dominant message at the embryonic stage (4-6 hrs), known to have zygotic gene expression (Fig. 4). No effort was made to isolate cDNA clones corresponding to the 4.7 kb transcript; thus the exact sequence of this short transcript is not known. However, a polyadenylation signal 35 consensus sequence was found at nucleotide position 4655 - 4660 in the 5.7 kb transcript and in the corresponding genomic DNA (Fig. 5) and a 0.51 kb probe from the 3' end of

the 5.7 kb transcript did not hybridize to the 4.7 kb transcript while a 1 kb probe from the 5' untranslated region of the 5.7 kb transcript hybridized to both the 5.7 kb and 4.7 kb transcripts. This suggests that the 4.7 kb transcript may be a truncated version of the 5.7 kb transcript. The genomic and cDNA sequence corresponding to the 5.7 kb transcript was determined (Materials and Methods). The entire 5720 bp cDNA sequence, which is interrupted by seven introns, and the putative lats product (lats), deduced from the long open reading frame, are illustrated in Figure 5. An interesting feature of the 5.7 kb transcript is the existence of a 141 bp segment located in the 3' untranslated region (Fig. 5), which is identical to the first 141 bp of the 5' untranslated region of the class I transcript from the *Drosophila* phospholipase C gene, *plc-21* (Shortridge et al., 1991, J. Biol. Chem. 266:12474-12480). The functional significance of this sequence motif is unknown. It could be a regulatory target sequence that is shared by both genes.

There are 34 differences between the lats cDNA and genomic sequences and 31 of them do not affect the deduced amino acid sequence. In the remaining three differences, one changes the serine 206 in cDNA into a cysteine. The second change in the genomic sequence adds an additional glutamine in the poly-glutamine opa repeat (Fig. 6; Wharton et al., 1985, Cell 40:55-62). The third is the addition of a fifteen bp sequence in the genomic DNA after the nucleotide 2644 of the cDNA. This sequence could be translated into another copy of the Arg-Glu-Arg-Asp-Gln (part of SEQ ID NO:2) peptide. However, this sequence is not present in the two independent cDNA clones that were sequenced.

The predicted lats product contains 1099 amino acid residues. The kinase domain of lats is more similar to protein-serine/threonine kinases than to protein-tyrosine kinases, especially in the sequences of the domains VI and VIII defined by Hanks et al. (1988, Science 241:42-52); protein-serine/threonine kinase consensus in domain VI: Asp-Leu-Lys-Pro-Glu-Asn (SEQ ID NO:9). Lats sequence in domain

VI: Arg-Asp-Ile-Lys-Pro-Asp-Asn (836-842) (part of  
SEQ ID NO:2); protein-serine/threonine kinase consensus in  
domain VIII: Gly-Thr/Ser-X-X-Tyr/Phe-X-Ala-Pro-Glu  
(SEQ ID NO:10). Lats sequenc in domain VIII: Gly-Thr-Pro-  
5 Asn-Tyr-Ile-Ala-Pro-Glu (917-925) (part of SEQ ID NO:2). The  
C-terminal half of lats shares extensive sequence similarity  
with a group of six proteins including the Dbf20 and Dbf2  
cell cycle protein-ser/thr kinases from *Saccharomyces*  
*cerevisiae* (Johnston et al., 1990, Mol. Cell. Biol. 10:1358-  
10 1366; Toyn et al., 1991, Gene 104:63-70; Toyn and Johnston,  
1994, EMBO J. 13:1103-1113), and the COT-1 putative protein  
kinase from *Neurospora crassa* (Yarden et al., 1992; EMBO J.  
11:2159-2166) (Fig. 6A, 6B). The sequence similarity between  
the kinase domains of lats and these proteins (39-49%  
15 identity) is much higher than the sequence similarity  
observed between the different subgroups of protein-ser/thr  
kinases (20-25% identity; Hanks et al., 1988, Science  
241:42-52). However, there is an insertion of about 40 amino  
acid residues within the kinase domains of these proteins,  
20 sharing little sequence similarity (denoted by a black bar in  
Fig. 6B). The human myotonic dystrophy protein kinases  
(MDPK) also have significant similarity with the C-terminal  
region of lats (Brook et al., 1992, Cell 68:799-808; Fu et  
al., 1993; Science 260:235-238, Mahadevan et al., 1993, Hum.  
25 Mol. Genet. 2:299-304), but their kinase domains do not  
contain this ~40 amino acid insertion. In addition, lats and  
these proteins also share significant levels of sequence  
similarity in the two regions (each contains ~100-150 amino  
acids) flanking the kinase domain (20-28% identity; Fig. 6A,  
30 6B). In the case of Dbf20, its entire sequence except for the  
20 C-terminal most residues can be aligned with lats,  
indicating lats is a close relative of Dbf20. A poly-  
glutamine opa repeat is located near the middle of the  
protein (Fig. 5; Wharton et al., 1985, Cell 40:55-62). The  
35 N-terminal half of lats contains many short homopolymeric  
runs including poly-prolin which makes up about 15% of the  
residues. At least one of the proline-rich stretches closely

match s the consensus of SH3-binding sites (Fig. 3B; Ren et al., 1993, Science 259:1157-1161), raising the possibility that it may interact with SH3-containing proteins. No putativ signal sequ nce app ars in the lats protein, 5 indicating that it is an intracellular protein.

### 6.3. DISCUSSION

#### Screening for Mutations in Mosaic Animals to Identify and Study Potential Tumor Suppressors

10 The comparison between mosaic flies and tumor patients is simplistic yet useful. Tumor patients contain wt tumor suppressor genes in most of their cells and only small groups of cells sustain mutations in tumor suppressors. We have searched for recessive overproliferation mutations in mosaic 15 animals. Flies that carry somatic cells mutated for tumor suppressors or negative regulators of cell proliferation are viable, yet the overproliferation mutant phenotype is readily identifiable. Therefore, mosaic flies, which are in a fashion analogous to tumor patients, provide a mean to screen 20 for potential tumor suppressors. Three overproliferation mutations were identified in our screen. They were not identified as "interesting" mutations in screens for embryonic lethal mutations. Identifying overproliferation mutations in homozygous mutant larvae and pupae is not only 25 biased against embryonic lethals, but also laborious, since it requires establishment of individual lines before examining the potential phenotypes. Further screens for overproliferation mutations in mosaic animals will allow us to identify other important players in pathways that 30 negatively regulate cell proliferation.

The overproliferation phenotypes that we observed were caused by loss of function in a single gene. In humans, it was suggested that most retinoblastomas are caused by defects in a single tumor suppressor (Knudson, 1971, Proc. Natl. 35 Acad. Sci. USA 68:820-823). On the other hand, evidence indicates that tumorigenesis in other human tissues (e.g.,



colon cancer) is a multistep process which involves inactivation of more than one gene (Fearon and Vogelstein, 1990, Cell 61:759-767; Vogelstein and Kinzler, 1993, Trends Genet. 9:138-141). Overproliferation caused by defects in multiple genes is unlikely to be detected in our screens unless these genes are located on the same chromosome arm. To identify this type of gene, one could perform a modified mosaic screen which induces clones of cells to become homozygous for more than one mutagenized chromosome arm.

10

#### lats Affects Many Tissues Throughout Development

The *lats* gene is genetically defined by a single complementation group that consists of various alleles causing a wide range of defects. Different alleles were found to cause lethality at almost every stage during development: embryo, early larvae, late larvae, early pupae, late pupae and pharate-adult. The embryonic lethality occurs in the pharate first instar stage. The early embryonic requirements for *lats* could well be masked by the wt products that are maternally deposited in the egg. Weak *lats* alleles produce viable animals with phenotypes ranging from rough eye to sterility. The *lats* transcripts were detected throughout development up to adult stage, consistent with the observation that *lats* mutants affect all these stages. Although mutations at *lats* cause many defects, affecting cell proliferation could cause most of the phenotypes including overproliferation in mutant clones, lethality at the various stages, tissue overproliferation on the head, broadened wing blade, and sterility in homozygous mutants. However, phenotypes such as extra cuticle deposits and malformed bristles and hairs are evidence of defects in differentiation.

The different behavior of the *lats* mutant clones and clones mutant for other previously identified *Drosophila* tumor suppressors is interesting. Cells mutant for *dlg*, *lgl* or *hyd* seem to fail to receive growth regulation signals. They proliferated slower than wt cells during larval stages

when the cells were instructed to proliferate, and they failed to terminate proliferation in late larval and pupal stages when the wt cells have ceased proliferation. On the other hand, the *lats* mutant clones induced during the larval stages were overproliferated, and later the mutant cells on the body were differentiated to form adult cuticular structures. Thus, *lats* could be a negative regulator that monitors the rate of proliferation.

The *lats* gene is located in a complex region. The 5' end of the *lats* 5.7 kb transcript (cDNA) is only about 550 bp away from the T2 transcript and its 3' end is about 1.5 kb away from the *zfh-1* transcript. Furthermore, all three of these closely located transcripts are located in an intron of the T1 transcription unit. Thus, a sizable deletion in the 5.7 kb transcription unit could affect the function of any of the genes in the region, which makes it difficult to determine which transcript is responsible for the *lats* phenotype. The fact that P-element transform lines carrying a cDNA from the 5.7 kb transcript under the *hsp70* promoter rescued all types of *lats* alleles demonstrated that the 5.7 kb transcription unit is the *lats* gene.

The *lats* Putative Protein-Ser/Thr Kinase  
Shares Homology With Proteins That Are  
Involved in Regulation of Cell Cycle  
and Growth in Budding Yeast and *Neurospora*

All 11 subdomains of the kinase domain that are found in previously identified protein kinases (Hanks et al., 1988, Science 241:42-52) are conserved in *lats*. This predicts that *lats* is a protein kinase. Furthermore, the sequence comparisons suggest *lats* to be a ser/thr kinase as the *lats* kinase domain is more similar to protein-ser/thr kinases than to protein-tyr kinases. The C-terminal half of *lats* shares extensive sequence similarity with a group of six proteins. Mutations are known for three of these genes and in each case they affect either cell cycle or growth. The *cot-1* (colonial temperature sensitive-1) gene of *Neurospora* was identified by a temperature sensitive mutant that causes compact colony

growth (Mitchell and Mitchell, 1954, Proc. Natl. Acad. Sci. USA 40:436-440; Galsworthy, 1966, Diss. Abstr. 26:6348). Wild-type filamentous ascomycete *Neurospora* grows on solid media by continuous hyphal elongation and branching to form

5 spreading colonies. Strains lacking functional *cot-1* gene are viable, but their hyphae branch extensively, resulting in compact colonial growth (Yarden et al., 1992, EMBO J. 11:2159-2166). This extensive branching phenotype is somewhat similar to the growth property of the *lats* mutant

10 clones: the *lats* mutant cells continue to "bud" out of the surface to form new proliferation lobes. Another homologous gene, the *DBF2* gene of the budding yeast, was identified in a genetic screen for mutations causing defects in DNA synthesis (Johnston and Thomas, 1982, Mol. Gen. Genet. 186:439-444).

15 The temperature sensitive alleles of *DBF2* were found to both delay the initiation of S phase and also to arrest the cell cycle during nuclear division (Johnston et al., 1990, Mol. Cell. Biol. 10:1358-1366). The *DBF20* gene was identified through cross hybridization with *DBF2* DNA (Toyn et al., 1991,

20 Gene 104:63-70). Strains carrying deletions for either *DBF2* or *DBF20* are viable; however, deleting both genes in the same strain causes lethality. The kinase activities of both proteins have been shown to be specific for serine/threonine residues and are regulated during the cell cycle (Toyn and

25 Johnston, 1994, EMBO J. 13:1103-1113). In the case of *Dbf20*, its entire sequence except the 20 most C-terminal residues can be aligned with *lats*. The mutant phenotype of *lats* and its sequence homology with the cell cycle protein kinases is consistent with the notion that *lats* might be directly

30 involved in regulation of the cell cycle. The N-terminal half of *lats* contains many proline-rich stretches and at least one of them closely matches the consensus sequence of SH3 binding sites (Ren et al., 1993, Science 259:1157-1161), raising the possibility that this region could be a

35 regulatory domain for the *lats* kinase, which binds to SH3 domain-containing proteins.

In recent years, many protein kinases have been identified to be involved in regulation of the cell cycle and cell proliferation. While Wee1 is an inhibitor of the Cdc2 kinase (Russell and Nurse, 1987, Cell 49:559-567; Featherstone and Russell, 1991, Nature 349:808-811), all other previously identified protein kinases are positive regulators of cell proliferation. They are either required for completion of the cell cycle or for signalling cells to proliferate. Lats is the first predicted protein-ser/thr kinase that has been shown to cause overproliferation when its function is removed. Studies of lats and other overproliferation mutations in *Drosophila* will provide a better understanding of how cell proliferation is regulated during development and how mutations could lead to abnormal growth.

#### 7. ISOLATION AND CHARACTERIZATION OF MAMMALIAN LATS HOMOLOGS

As described herein, we have cloned and sequenced both mouse and human lats homologs.

##### 7.1. ISOLATION AND CHARACTERIZATION OF MOUSE LATS HOMOLOGS

cDNA clones for two different lats homologs in mice were obtained as follows.

##### 25 Screening of Mouse Homologs:

Probe: A 2.2 kb BamHI fragment containing the kinase domain of the *Drosophila* lats gene was labeled with <sup>32</sup>P by random labeling

Library: Newborn mouse brain lambda ZAP cDNA library from Stratagene

##### 30 Hybridization

Condition:	45°C, overnight in	6x	SSC
		5x	Denhart's
		0.5%	SDS (sodium dodecyl sulfate)
		100 µg/ml	salmon sperm DNA

35 Wash:	50°C, 30 min. x 4, in	2x	SSC
		0.1%	SDS

**Results:** Three positive clones were identified. (M41 clone for the *m-lats* gene, and M51 and M31 clones for the *m-lats2* gene.)

Two different mouse *lats* homologs, termed *m-lats* and *m-lats2*, respectively, were isolated and sequenced. Both the *m-lats* and *m-lats2* clones are missing a small amount of the 5' end of their respective genes. The cDNA sequence (SEQ ID NO:5) and deduced protein sequence (SEQ ID NO:6) of *m-lats* are shown in Figure 7. The cDNA sequence (SEQ ID NO:7) and deduced protein sequence (SEQ ID NO:8) of *m-lats2* are shown in Figure 8.

Portions of both the *m-lats* and *m-lats2* cDNAs were used as probes to screen a mouse genomic library, under standard hybridization conditions. Genomic clones for both *m-lats* and *m-lats2* have been isolated that contain most of the coding regions of these genes.

## 7.2. ISOLATION AND CHARACTERIZATION OF HUMAN *LATS* HOMOLOGS

cDNA clones for at least one human *lats* homolog were obtained as follows.

Screening of Human Homologs (moderately stringent conditions):

**Probe:** A 2.1 kb PstI fragment containing the kinase domain of the *m-lats* gene was labeled with <sup>32</sup>P by random labeling

**Library:** Fetal human brain lambda gt10 cDNA library from Clontech

**Hybridization Condition:** 55°C, overnight in 6x 5x 0.5% 100 µg/ml SSC Denhart's SDS salmon sperm DNA

**Wash:** 60°C, 30 min. x 2, in 1x 0.1% SSC SDS

**Results:** About 20 positive clones were identified for the *h-lats* gene.

One human *lats* homolog, termed *h-lats*, was isolated and sequenced. The cDNA sequence (SEQ ID NO:3) and deduced

protein s quence (SEQ ID NO:4) of *h-lats* are shown in Figur  
9. Th d duced protein sequence is full-length. The  
c mpl te coding sequence of the *h-lats* cDNA was inserted into  
a bact rial cloning vector (derived from Bluescript (KS)-  
5 vector; Stratagene) to form plasmid pBS(KS)-*h-lats* (Fig. 10).  
The total size of pBS(KS)-*h-lats* is 6.96 kb.

A *h-lats* cDNA fragment was used as a probe under  
conditions of moderate stringency to screen a human genomic  
cosmid library. Genomic *h-lats* clones were isolated. Over  
10 70 kb of the genomic *h-lats* sequence has been isolated; the  
isolated sequences include all of the *h-lats* coding sequence  
(but not all the exon sequences).

An *m-lats2* cDNA fragment was used as a probe to  
screen a human genomic phage library under the conditions  
15 described above, except that hybridization was carried out at  
50°C, and washing was carried out at 55°C with 2X SSC, 0.1%  
SDS. Two genomic *h-lats* clones have been isolated that  
specifically hybridize to *m-lats2* cDNA probes and do not  
hybridize to *m-lats* and *h-lats* cDNA probes.

20

#### 8. CONSERVATION OF SEQUENCES AND DOMAIN STRUCTURE AMONG LATS HOMOLOGS OF DIFFERENT SPECIES

Comparison of the sequences of *Drosophila* lats,  
*h-lats*, *m-lats*, and *m-lats2* showed a startlingly high degree  
25 of sequence conservation, both overall and within domains of  
the lats protein. An alignment of the *h-lats* (SEQ ID NO:4)  
and *m-lats* (SEQ ID NO:6) protein sequences is shown in Figure  
11. The overall amino acid sequence identity between *h-lats*  
and *m-lats* is 93%. An alignment of the *h-lats* (SEQ ID NO:4)  
30 and *m-lats2* (SEQ ID NO:8) protein sequences is shown in  
Figure 12.

Homologous domains (*i.e.*, domains conserved)  
between the different lats homologs were identified. Figure  
13 presents an alignment of the *h-lats* protein sequence (SEQ  
35 ID NO: 4) and the *Drosophila* lats protein sequence (SEQ ID  
NO:2), and indicates the domains identified as conserved  
among the lats homologs from th various species.

The identified domains were as follows:

(1) Lats C-terminal domain 3 (LCD3)

The last three amino acids (VYV) are completely conserved in all four homologs including *Drosophila* lats, h-lats, m-lats, and m-lats2.

(2) Lats C-terminal domain 2 (LCD2)

	amino acid residues
h-lats	1077-1086
<i>Drosophila</i> lats	1075-1084

This domain is completely conserved in all four homologs including *Drosophila* lats, h-lats, m-lats, and m-lats2 (10/10 identical residues).

(3) Lats C-terminal domain 1 (LCD1)

	amino acid residues
h-lats	1032-1043
<i>Drosophila</i> lats	1035-1047

This domain is completely conserved among *Drosophila* lats, h-lats, and m-lats (12/12 identical), and is highly conserved between any of the foregoing and m-lats2 (11/12 identical).

(4) Kinase domain

	amino acid residues
h-lats	703-1014
<i>Drosophila</i> lats	711-1018

This domain is highly conserved among the four homologs (76% identical between *Drosophila* lats and h-lats; 99% identical between h-lats and m-lats; 83% identical between h-lats and m-lats2).

A potential phosphorylation residue in *Drosophila* lats and the mammalian homologs that could lead to the activation of the lats kinases after phosphorylation was identified.

Activities of protein kinases are often regulated by varying the phosphorylation state of specific serine, threonine, and tyrosine residues. Phosphorylation of a serine or threonine within twenty residues upstream of

an Ala-Pro-Glu consensus in subdomain eight of the kinase domain, is often required for catalytic activities of many protein-ser/thr kinases (Hanks et al., 1988, Science 241:42-52). For example, Thr167 and Thr197 are phosphorylated in Cdc2 of fission yeast and in the cardiac muscle adenosine 3',5'-phosphate dependent protein kinase, respectively (Ducommun et al., 1991, EMBO J. 10:3311-3319; Gould et al., 1991, EMBO J. 10:3297-3309; Shoji et al., 1983, Biochem.

22:3702-3709). A ser residue in a similar position of the lats kinase domain is conserved in *Drosophila* lats, h-lats, m-lats, and m-lats2 (Ser914 in *Drosophila* lats; Ser909 in h-lats). Thus, the activities of *Drosophila* lats and its mammalian homologs may be regulated by phosphorylation of this ser residue.

(5) Lats flanking domain (LFD)

	amino acid residues
h-lats	607-702
<i>Drosophila</i> lats	612-710

LFD is a domain that flanks and is amino-terminal to the kinase domain. This domain is highly conserved between *Drosophila* lats and h-lats (68% identical) and is also highly conserved between h-lats and m-lats2 (71% identical). This domain is completely conserved between h-lats and m-lats (100% identical).

(6) Lats split domain 1 (LSD1)

	amino acid residues
LSD1 <i>Drosophila</i> -lats	365-392
LSD1 anterior (LSD1a) h-lats	328-334
LSD1 posterior (LSD1p) h-lats	498-518

Certain lats domains have been termed split domains because the amino- (anterior) and carboxy- (posterior) portions of the domain appear separated from each other in at least one of the lats homologs. Split domains may constitute discontinuous binding/functional regions (e.g., brought together by tertiary structure). The LSD1a subdomain is completely conserved among *Drosophila*



lats, h-lats, and m-lats (7/7 identical), and is not conserved in m-lats. The LSD1p subdomain is conserved between the four homologs (14/21 identical among *Drosophila* lats, h-lats, and m-lats; 13/21 identical between h-lats and m-lats<sup>2</sup>). The LSD1a and LSD1p subdomains are adjacent to each other in *Drosophila* lats and are separated in the mammalian homologs.

(7) Lats split domain 2 (LSD2)

		amino acid residues
10	LSD2 <i>Drosophila</i> lats	536-544
	LSD2 anterior (LSD2a) h-lats	28-31
	LSD2 posterior (LSD2p) h-lats	555-559

Both the LSD2a and LSD2p subdomains are completely conserved among the four homologs. However, the two subdomains are adjacent to each other in *Drosophila* lats and are separated in the mammalian homologs.

(8) Putative SH3-binding domain (SH3-binding)

		amino acid residues
	h-lats	247-268
	<i>Drosophila</i> lats	196-217

This domain is highly conserved among *Drosophila* lats, h-lats, and m-lats (10/22 identical), and does not exist in m-lats<sup>2</sup>.

The opa domain does not appear in the mammalian lats homologs.

9. FUNCTIONAL INTERCHANGEABILITY OF THE HUMAN AND *DROSOPHILA* LATS HOMOLOGS

9.1. OVEREXPRESSION OF HUMAN LATS OR OF *DROSOPHILA* LATS CAUSES A SMALLER, ROUGH EYE IN *DROSOPHILA*

Overexpression of lats and h-lats in the developing *Drosophila* eye was carried out. The *Drosophila* lats cDNA and h-lats cDNA were each cloned into the pGMR P-element vector. This vector was constructed by Bruce Hay and Gerald M. Rubin at the University of California at Berkeley, and will direct the expression of a cDNA of interest in the posterior region of the developing third instar larval eye imaging disc of

*Drosophila*. Ten independent transformant lines for each of the pGMR-lats and pGMR-h-lats constructs were generated. The adult eyes of all these lines displayed a small-rough eye phenotype (eyes smaller than normal, with irregular, rough appearance). This indicates that both lats and h-lats genes have the same biological effect when they are overexpressed in the developing *Drosophila* eye.

10 9.2. HUMAN H-LATS GENE CAN REPLACE THE  
DROSOPHILA HOMOLOG TO PREVENT  
DEATH IN DROSOPHILA ANIMALS HAVING  
MUTANT DROSOPHILA LATS

The *Drosophila* lats cDNA was cloned into the pCaSpeR-hs vector (Thummel and Pirrotta, 1992, *Drosophila Inform. Service* 71:150) for germ line transformation of *Drosophila*. Three of the transformed lines were tested and were able to rescue the lethality of the *lats<sup>al</sup>/lats<sup>xl</sup>*, *lats<sup>Pl</sup>* and *lats<sup>c26-1</sup>* animals after one hour heat shock for every 24 hours during larval and pupal development. The human h-lats cDNA (in a XhoI (blunted)-XbaI fragment) from pBS(SK)-h-lats (Fig. 10) was cloned into the HpaI-XbaI sites of the pCaSpeR-hs vector, to produce plasmid pCaSpeR-hs-h-lats (Fig. 14). Plasmid pCaSpeR-hs-h-lats was used for germ line transformant. Three of the pCaSpeR-hs-h-lats transformant lines were tested and were able to rescue the lethality of the *lats<sup>Pl</sup>* and *lats<sup>c26-1</sup>* animals under the same conditions used in rescuing experiments for the *Drosophila* gene.

30 10. HUMAN LATS EXPRESSION IS FOUND IN ALL  
NORMAL TISSUES TESTED AND IS ABSENT  
IN A LARGE NUMBER OF TUMOR CELL LINES

10.1. HUMAN LATS EXPRESSION IN NORMAL TISSUES

The expression of human lats RNA was investigated in various adult tissues. A 1.2 kb BamHI fragment of the h-lats cDNA was used as a <sup>32</sup>P-labeled probe for Northern analysis. Hybridization was to a nylon membrane containing polyA<sup>+</sup> RNA from various human fetal and adult tissues, obtained from Clontech. The Northern analysis was carried

out according to the recommended instructions of the manufacturer (Clontech). The results are shown in Figure 15. h-lats was expressed in every tissue tested (fetal brain, fetal lung, fetal liver, fetal kidney, adult spleen, adult thymus, adult prostate, adult testis, adult ovary, adult small intestine, adult colon, and adult blood leukocytes). Expression was higher in fetal tissues than in adult tissues.

#### 10.2. HUMAN LATS EXPRESSION IN VARIOUS TUMOR CELL LINES

10 The <sup>32</sup>P-labeled BamHI fragment of h-lats was used as a probe for Northern analysis, for hybridization to total RNAs isolated from 42 different human tumor cell lines (obtained from the American Type Culture Collection, Rockville, MD). No h-lats expression was detected in 20 of the tumor lines (48%). The name and tissue origin of the tumor cell lines tested, and the results of the Northern analysis are presented in Table 3.

20

Table 3

	<u>Name of tumor lines</u>	<u>Tumor Origin</u>	<u>Expression detected by Northern analyses</u>	
			<u>YES</u>	<u>NO</u>
25	5637	Bladder		X
	RT4	Bladder	±*	
	HT-1376	Bladder		X
	HT-1197	Bladder		X
30	BT-20	Breast	X	
	BT-474	Breast	X	
	ZR-75-1	Breast		X
	ZR-75-30	Breast	X	
	BT-549	Breast		X
	MDA-MB-453	Breast		X
	MDA-MB-435S	Breast		X
	HBL-100	Breast		X
35	LoVo	Colon		X
	HT-29	Colon	X	
	HCT116	Colon	X	
	LS 180	Colon		X
	DLD-1	Colon	X	
	WiDr	Colon	X	

	SW480	Colon	X	
	Caco-2	Col n	±	
	HEL 92.1.7	Erythroleukemia	X	
	MOLT-4	Leuk mia	X	
	CEM-CM3	Leukemia	X	
5	K-562	Leukemia	X	
	Jurkat	Leukemia		X
	HUT 78	Lymphoma	X	
	SK-LU-1	Lung		X
	A-427	Lung		X
	Calu-1	Lung	X	
10	NCI-H69	Lung	X	
	SK-MEL-3	Melanoma		X
	SK-MEL-28	Melanoma		X
	SK-MEL-31	Melanoma		X
	MIA PaCa-2	Pancreas		X
	BxPC-3	Pancreas		X
15	Hs 700T	Pancreas	X	
	Hs 766T	Pancreas	X	
	RD	Sarcoma		X
	A-204	Sarcoma		X
	AN3 CA	Uterine	X	
20	SK-UT-1	Uterine	X	
	HEC-1-A	Uterine	±	

\*: weak signal

25 Thus, 48% of the tumor cell lines tested had no detectable *h-lats* expression, whereas 100% of the normal tissues tested had detectable *h-lats* expression. It should be noted that the 48% figure may be an underestimate of the actual number of tumor cell lines that had decreased *lats* protein level or activity relative to normal tissue, since 30 while lack of *lats* RNA (i.e., a transcriptional block) allows the conclusion that no *lats* protein is made, tumor cells that expressed the *lats* RNA may still have had no or low *lats* protein levels and/or activity due to the possible existence 35 of a translational block or the presence of mutation(s) in an expressed *lats* protein.

# 11. DEPOSIT OF MICROORGANISM

Bacteria strain *E. coli* TG2 containing plasmid pBS(KS)-h-lats was deposited on March 24, 1995 with the American Type Culture Collection, 1201 Parklawn Drive, 5 Rockville, Maryland 20852, under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedures, and assigned Accession No. 69769.

10           The present invention is not to be limited in scope by the microorganism deposited or the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing  
15 description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Various references are cited herein, the disclosures of which are incorporated by reference in their entireties.

20

25

30

35

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Xu, Tian  
Tao, Wufan  
Wang, Weyi  
Zhang, Sheng  
Yu, Wan
- (ii) TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS  
GENES AND METHODS BASED THEREON
- (iii) NUMBER OF SEQUENCES: 16
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Pennie & Edmonds
  - (B) STREET: 1155 Avenue of the Americas
  - (C) CITY: New York
  - (D) STATE: New York
  - (E) COUNTRY: USA
  - (F) ZIP: 10036-2711
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: To Be Assigned
  - (B) FILING DATE: On Even Date Herewith
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Misrock, S. Leslie
  - (B) REGISTRATION NUMBER: 18,872
  - (C) REFERENCE/DOCKET NUMBER: 6523-007
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (212) 790-9090
  - (B) TELEFAX: (212) 869-9741/8864
  - (C) TELEX: 66141 PENNIE

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5720 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1103..4402
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATCTAGCAGC ACGGCAGCAA CAAAACCACG AATTAATTTT ACTAAATTTA AGCCAAACGC

60

GCATCGGAAA TGCCTGAAAA TGCATTGAA TGCACGCGAA AAGTGATGGG TTGCGAACGC	120
GA TGAATCA AGTGAAAATA CGTCGGCAAA TATCAGCGAA TTGTCGTCAA AAGGCAAGGA	180
AAAACGGAGA AAAAGAGGAA AAGCAATAAG TGCCGTGTGT GGGAAACGCG AAAAAGGCGA	240
GAACAAAGAG GCGAAAAGCG AGGAAATTGC GTGGAAAACG TGGAAAACGC GAAGAAGCGA	300
AGCTCCAAGT TGGCCGCCAT CGATTCTGTC GTAGGATCAA TTAAGATTCC GAGTGGTCTGA	360
GAATCGGCTC AAATCAAATT AAAATCAACT AATATTTTGG TATTCAGATA TTCAAATGGA	420
ATTCATTCAT CGCCTGCGAC TTTTATTCTGG ATCTGCCAAC TATTTTGTAA TTTGAATTGT	480
GTGTCTGCGG CTGGCGCAGA ATCTCTGATA AAGCAGAGGA ATAAAATCGG AAGAACAACA	540
AATACAAATA CAAATGAAAT GCGGGGAGCA GTATTTACAT GCCAAATGAA TGCTGGATAG	600
GCGAAAGGGG GGGTTTCTCT TATAATGCAA ATGTGAATGT GAATGCGAAT GCGAATGCGA	660
GTGGAAGAAT TCCCGGCGCG AGTGATAAAT AATCCGACGA CAAACAAAGC AGAAGCCTAC	720
ACCGCGAGAA AGAGCAGCGC AAACACAATT ATCTTTATTG AGAGCAACAA TATCAAGATC	780
GAGATAATAA AGCATCCTAA AACCCGCGCC TTAGTTCGTT TTAGTCTCGC CACGGATATA	840
GATATTCAAA GGCAAAAAGG TGGTGTGCGC ATCGCCAGAC AAACAAGTAA AGCATCTATT	900
TCATACAAAA CAACCAATTA AATAATAATA AAAATAATAA TAATCGTAGA GAGGCAGAGC	960
CAAATCAAAT TCCCGGCCGC CGATGTGCCC CAGTGTGTGT GCGTGTGTGT GTGTGTGTGC	1020
TGTGCTGTGC TGTGCGAGTG TTAGTGTGCG GAGCATTTCT GTGATATGAG TGCTAAATGC	1080
CACAGGGCGA AGCAGCAGCA TC ATG CAT CCA GCG GGC GAA AAA AGG GGC GGT	1132
Met His Pro Ala Gly Glu Lys Arg Gly Gly	10
CGC CCC AAT GAT AAA TAC ACG GCG GAA GCC CTC GAG AGC ATC AAG CAG	1180
Arg Pro Asn Asp Lys Tyr Thr Ala Glu Ala Leu Glu Ser Ile Lys Gln	25
GAC CTA ACC CGA TTT GAA GTA CAA AAT AAC CAT AGG AAT AAT CAG AAT	1228
Asp Leu Thr Arg Phe Glu Val Gln Asn Asn His Arg Asn Asn Gln Asn	40
TAC ACA CCT CTG CGA TAC ACG GCG ACC AAC GGA CGC AAC GAT GCA CTT	1276
Tyr Thr Pro Leu Arg Tyr Thr Ala Thr Asn Gly Arg Asn Asp Ala Leu	55
ACT CCT GAC TAT CAC CAC GCC AAG CAG CCG ATG GAG CCG CCA CCC TCC	1324
Thr Pro Asp Tyr His His Ala Lys Gln Pro Met Glu Pro Pro Pro Ser	70
GCC TCT CCT GCT CCG GAC GTG GTC ATA CCG CCG CCG CCC GCC ATT GTA	1372
Ala Ser Pro Ala Pro Asp Val Val Ile Pro Pro Pro Pro Ala Ile Val	90
GGT CAG CCC GGA GCC GGC TCC ATA TCC GTA TCC GGT GTG GGC GTT GGA	1420
Gly Gln Pro Gly Ala Gly Ser Ile Ser Val Ser Gly Val Gly Val Gly	105
GTG GTG GGT GTG GCG AAC GGA CGT GTG CCA AAG ATG ATG ACG GCC CTA	1468
Val Val Gly Val Ala Asn Gly Arg Val Pro Lys Met Met Thr Ala Leu	120
ATG CCA AAC AAA CTG ATC CGG AAG CCG AGC ATC GAA CGG GAC ACG GCG	1516

Met	Pro	Asn	Lys	Leu	Ile	Arg	Lys	Pro	Ser	Ile	Glu	Arg	Asp	Thr	Ala	
		125					130					135				
AGC	AGT	CAC	TAC	CTG	CGC	TGC	AGT	CCG	GCT	CTG	GAC	TCC	GGA	GCC	GGT	1564
Ser	Ser	His	Tyr	Leu	Arg	Cys	Ser	Pro	Ala	Leu	Asp	S r	Gly	Ala	Gly	
		140				145					150					
AGC	TCC	CGA	TCG	GAC	AGC	CCC	CAT	TCG	CAC	CAC	ACC	CAC	CAG	CCG	AGC	1612
Ser	Ser	Arg	Ser	Asp	Ser	Pro	His	Ser	His	His	Thr	His	Gln	Pro	Ser	
		155			160					165					170	
TCG	AGG	ACG	GTG	GGT	AAT	CCA	GGT	GGA	AAT	GGT	GGA	TTT	TCT	CCG	TCG	1660
Ser	Arg	Thr	Val	Gly	Asn	Pro	Gly	Gly	Asn	Gly	Gly	Phe	Ser	Pro	Ser	
				175					180					185		
CCA	AGC	GGT	TTC	AGT	GAG	GTG	GCT	CCA	CCG	GCG	CCG	CCG	CCA	CGC	AAT	1708
Pro	Ser	Gly	Phe	Ser	Glu	Val	Ala	Pro	Pro	Ala	Pro	Pro	Pro	Arg	Asn	
			190					195					200			
CCC	ACC	GCC	TCC	AGC	GCG	GCC	ACG	CCC	CCA	CCG	CCA	GTG	CCG	CCC	ACC	1756
Pro	Thr	Ala	Ser	Ser	Ala	Ala	Thr	Pro	Pro	Pro	Pro	Val	Pro	Pro	Thr	
		205					210					215				
AGC	CAG	GCG	TAC	GTG	AAG	CGG	CGA	TCA	CCG	GCC	CTG	AAC	AAC	CGC	CCG	1804
Ser	Gln	Ala	Tyr	Val	Lys	Arg	Arg	Ser	Pro	Ala	Leu	Asn	Asn	Arg	Pro	
		220				225					230					
CCG	GCG	ATA	GCG	CCA	CCC	ACT	CAG	CGA	GGC	AAC	TCA	CCT	GTA	ATA	ACC	1852
Pro	Ala	Ile	Ala	Pro	Pro	Thr	Gln	Arg	Gly	Asn	Ser	Pro	Val	Ile	Thr	
		235			240					245					250	
CAA	AAC	GGG	CTG	AAG	AAC	CCG	CAG	CAG	CAG	TTG	ACG	CAG	CAG	CTG	AAG	1900
Gln	Asn	Gly	Leu	Lys	Asn	Pro	Gln	Gln	Gln	Leu	Thr	Gln	Gln	Leu	Lys	
				255					260					265		
TCC	CTG	AAC	CTA	TAC	CCA	GGC	GGA	GGC	AGT	GGA	GCA	GTG	GTG	GAG	CCA	1948
Ser	Leu	Asn	Leu	Tyr	Pro	Gly	Gly	Gly	Ser	Gly	Ala	Val	Val	Glu	Pro	
			270					275					280			
CCG	CCG	CCC	TAC	CTA	ATT	CAA	GGC	GGA	GCC	GGA	GGA	GCA	GCA	CCG	CCG	1996
Pro	Pro	Pro	Tyr	Leu	Ile	Gln	Gly	Gly	Ala	Gly	Gly	Ala	Ala	Pro	Pro	
		285					290					295				
CCG	CCA	CCA	CCC	AGT	TAC	ACG	GCC	TCC	ATG	CAG	TCG	CGG	CAG	TCG	CCC	2044
Pro	Pro	Pro	Pro	Ser	Tyr	Thr	Ala	Ser	Met	Gln	Ser	Arg	Gln	Ser	Pro	
		300				305					310					
ACA	CAA	TCC	CAA	CAA	TCG	GAC	TAC	AGG	AAA	TCC	CCG	AGC	AGT	GGG	ATA	2092
Thr	Gln	Ser	Gln	Gln	Ser	Asp	Tyr	Arg	Lys	Ser	Pro	Ser	Ser	Gly	Ile	
					320					325					330	
TAC	TCG	GCC	ACC	TCG	GCG	GGC	TCG	CCG	AGC	CCC	ATA	ACT	GTG	TCG	CTG	2140
Tyr	Ser	Ala	Thr	Ser	Ala	Gly	Ser	Pro	Ser	Pro	Ile	Thr	Val	Ser	Leu	
				335					340					345		
CCG	CCG	GCG	CCG	CTG	GCG	AAG	CCA	CAA	CCA	CGA	GTC	TAC	CAG	GCC	AGG	2188
Pro	Pro	Ala	Pro	Leu	Ala	Lys	Pro	Gln	Pro	Arg	Val	Tyr	Gln	Ala	Arg	
			350					355					360			
AGT	CAG	CAG	CCG	ATC	ATC	ATG	CAG	AGT	GTG	AAG	AGC	ACG	CAG	GTC	CAA	2236
Ser	Gln	Gln	Pro	Ile	Ile	Met	Gln	Ser	Val	Lys	Ser	Thr	Gln	Val	Gln	
		365					370					375				
AAG	CCC	GTG	CTG	CAA	ACA	GCA	GTG	GCG	CGC	CAA	TCG	CCA	TCG	AGT	GCC	2284
Lys	Pro	Val	Leu	Gln	Thr	Ala	Val	Ala	Arg	Gln	Ser	Pro	Ser	Ser	Ala	
		380				385					390					



TCG	GCC	AGC	AAT	TCA	CCA	GTC	CAC	GTG	CTG	GCC	GCT	CCA	CCC	TCT	TAC	2332
Ser	Ala	Ser	Asn	Ser	Pr	Val	His	Val	Leu	Ala	Ala	Pro	Pro	Ser	Tyr	
395					400					405					410	
CCT	CAG	AAG	TCC	CG	GCA	GTG	GTG	CAG	CAG	CAG	CAA	CAG	GCA	GCA	CG	2380
Pro	Gln	Lys	Ser	Ala	Ala	Val	Val	Gln	Gln	Gln	Gln	Gln	Ala	Ala	Ala	
				415					420					425		
GCG	GCC	CAC	CAG	CAG	CAG	CAT	CAG	CAC	CAG	CAA	TCC	AAA	CCA	CCA	ACG	2428
Ala	Ala	His	Gln	Gln	Gln	His	Gln	His	Gln	Gln	Ser	Lys	Pro	Pro	Thr	
			430				435						440			
CCA	ACC	ACA	CCG	CCC	TTG	GTG	GGT	CTG	AAC	AGC	AAG	CCC	AAT	TGC	CTG	2476
Pro	Thr	Thr	Pro	Pro	Leu	Val	Gly	Leu	Asn	Ser	Lys	Pro	Asn	Cys	Leu	
		445					450					455				
GAG	CCA	CCG	TCC	TAT	GCC	AAG	AGC	ATG	CAG	GCC	AAG	GCG	GCC	ACG	GTG	2524
Glu	Pro	Pro	Ser	Tyr	Ala	Lys	Ser	Met	Gln	Ala	Lys	Ala	Ala	Thr	Val	
	460					465					470					
GTA	CAG	CAG	CAG	CAA	CAG	CAG	CAG	CAA	CAA	CAG	CAG	GTC	CAG	CAG	CAG	2572
Val	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Val	Gln	Gln	Gln	
475				480					485						490	
CAG	GTG	CAA	CAG	CAG	CAG	CAA	CAG	CAG	CAA	CAG	CAA	CTG	CAG	GCC	TTG	2620
Gln	Val	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Leu	Gln	Ala	Leu	
				495				500						505		
AGG	GTG	CTC	CAG	GCA	CAG	GCT	CAG	AGG	GAG	CGG	GAT	CAA	CGG	GAG	CGG	2668
Arg	Val	Leu	Gln	Ala	Gln	Ala	Gln	Arg	Glu	Arg	Asp	Gln	Arg	Glu	Arg	
			510				515						520			
GAA	CGG	GAT	CAG	CAG	AAG	CTG	GCC	AAC	GGA	AAT	CCT	GGC	CGG	CAG	ATG	2716
Glu	Arg	Asp	Gln	Gln	Lys	Leu	Ala	Asn	Gly	Asn	Pro	Gly	Arg	Gln	Met	
		525					530					535				
CTT	CCG	CCG	CCG	CCC	TAT	CAG	AGC	AAC	AAC	AAC	AAC	AAC	AGC	GAG	ATC	2764
Leu	Pro	Pro	Pro	Pro	Tyr	Gln	Ser	Asn	Asn	Asn	Asn	Asn	Ser	Glu	Ile	
	540					545						550				
AAA	CCG	CCG	AGC	TGC	AAC	AAC	AAC	AAC	ATA	CAG	ATA	AGC	AAC	AGC	AAC	2812
Lys	Pro	Pro	Ser	Cys	Asn	Asn	Asn	Asn	Ile	Gln	Ile	Ser	Asn	Ser	Asn	
555					560					565					570	
CTG	GCG	ACG	ACA	CCA	CCC	ATT	CCG	CCT	GCC	AAA	TAC	AAT	AAC	AAC	TCC	2860
Leu	Ala	Thr	Thr	Pro	Pro	Ile	Pro	Pro	Ala	Lys	Tyr	Asn	Asn	Asn	Ser	
				575					580					585		
TCC	AAC	ACG	GGC	GCG	AAT	AGC	TCG	GGC	GGC	AGC	AAC	GGA	TCC	ACC	GGC	2908
Ser	Asn	Thr	Gly	Ala	Asn	Ser	Ser	Gly	Gly	Ser	Asn	Gly	Ser	Thr	Gly	
			590					595					600			
ACC	ACC	GCC	TCC	TCG	TCG	ACC	AGC	TGC	AAG	AAG	ATC	AAG	CAC	GCC	TCG	2956
Thr	Thr	Ala	Ser	Ser	Ser	Thr	Ser	Cys	Lys	Lys	Ile	Lys	His	Ala	Ser	
		605					610					615				
CCC	ATC	CCG	GAG	CGC	AAG	AAG	ATC	TCC	AAG	GAG	AAG	GAG	GAG	GAG	CGC	3004
Pro	Ile	Pro	Glu	Arg	Lys	Lys	Ile	Ser	Lys	Glu	Lys	Glu	Glu	Glu	Arg	
	620				625					630						
AAG	GAG	TTC	CGC	ATC	AGG	CAG	TAC	TCG	CCG	CAA	GCC	TTC	AAG	TTC	TTC	3052
Lys	Glu	Phe	Arg	Ile	Arg	Gln	Tyr	Ser	Pro	Gln	Ala	Phe	Lys	Phe	Phe	
635				640						645					650	
ATG	GAG	CAG	CAC	ATA	GAG	AAC	GTG	ATC	AAG	TCG	TAT	CGC	CAG	CGC	ACG	3100
Met	Glu	Gln	His	Ile	Glu	Asn	Val	Ile	Lys	Ser	Tyr	Arg	Gln	Arg	Thr	
				655					660					665		

TAT	CGC	AAG	AAT	CAG	CTG	GAG	AAG	GAG	ATG	CAC	AAA	GTG	GGA	CTG	CCC	3148
Tyr	Arg	Lys	Asn	Gln	Leu	Glu	Lys	Glu	Met	His	Lys	Val	Gly	Leu	Pro	
			670					675					680			
GAT	CAG	ACC	CAA	ATC	GAG	ATG	AGG	AAA	ATG	CTG	AAC	CAA	AAG	GAG	AGC	3196
Asp	Gln	Thr	Gln	Ile	Glu	Met	Arg	Lys	Met	Leu	Asn	Gln	Lys	Glu	Ser	
		685					690					695				
AAC	TAC	ATT	CGA	TTG	AAG	CGC	GCC	AAG	ATG	GAC	AAG	AGC	ATG	TTC	GTC	3244
Asn	Tyr	Ile	Arg	Leu	Lys	Arg	Ala	Lys	Met	Asp	Lys	Ser	Met	Phe	Val	
	700					705					710					
AAA	CTG	AAG	CCC	ATT	GGA	GTG	GGT	GCA	TTT	GGC	GAG	GTA	ACG	CTG	GTG	3292
Lys	Leu	Lys	Pro	Ile	Gly	Val	Gly	Ala	Phe	Gly	Glu	Val	Thr	Leu	Val	
	715				720					725					730	
AGC	AAA	ATC	GAT	ACC	TCG	AAC	CAT	TTG	TAT	GCG	ATG	AAA	ACC	CTG	CGG	3340
Ser	Lys	Ile	Asp	Thr	Ser	Asn	His	Leu	Tyr	Ala	Met	Lys	Thr	Leu	Arg	
				735					740					745		
AAA	GCG	GAC	GTT	CTC	AAG	CGG	AAT	CAG	GTG	GCA	CAC	GTG	AAG	GCC	GAG	3388
Lys	Ala	Asp	Val	Leu	Lys	Arg	Asn	Gln	Val	Ala	His	Val	Lys	Ala	Glu	
			750					755					760			
AGG	GAT	ATC	CTC	GCG	GAA	GCC	GAC	AAT	AAC	TGG	GTG	GTG	AAG	TTG	TAC	3436
Arg	Asp	Ile	Leu	Ala	Glu	Ala	Asp	Asn	Asn	Trp	Val	Val	Lys	Leu	Tyr	
		765					770					775				
TAC	AGC	TTC	CAG	GAC	AAG	GAT	AAT	CTG	TAC	TTT	GTG	ATG	GAC	TAC	ATA	3484
Tyr	Ser	Phe	Gln	Asp	Lys	Asp	Asn	Leu	Tyr	Phe	Val	Met	Asp	Tyr	Ile	
	780					785					790					
CCA	GGT	GGT	GAT	CTG	ATG	TCG	CTG	CTC	ATC	AAA	CTG	GGC	ATT	TTC	GAG	3532
Pro	Gly	Gly	Asp	Leu	Met	Ser	Leu	Leu	Ile	Lys	Leu	Gly	Ile	Phe	Glu	
	795				800					805					810	
GAG	GAA	CTG	GCC	AGA	TTC	TAC	ATC	GCC	GAG	GTC	ACC	TGC	GCC	GTG	GAC	3580
Glu	Glu	Leu	Ala	Arg	Phe	Tyr	Ile	Ala	Glu	Val	Thr	Cys	Ala	Val	Asp	
				815					820					825		
AGC	GTT	CAC	AAA	ATG	GGC	TTC	ATT	CAC	AGA	GAC	ATC	AAG	CCT	GAC	AAC	3628
Ser	Val	His	Lys	Met	Gly	Phe	Ile	His	Arg	Asp	Ile	Lys	Pro	Asp	Asn	
			830					835					840			
ATA	CTC	ATC	GAT	AGG	GAC	GGA	CAC	ATA	AAG	CTC	ACC	GAC	TTT	GGC	CTG	3676
Ile	Leu	Ile	Asp	Arg	Asp	Gly	His	Ile	Lys	Leu	Thr	Asp	Phe	Gly	Leu	
		845				850						855				
TGC	ACG	GGA	TTC	CGA	TGG	ACG	CAC	AAC	TCG	AAG	TAC	TAC	CAG	GAG	AAC	3724
Cys	Thr	Gly	Phe	Arg	Trp	Thr	His	Asn	Ser	Lys	Tyr	Tyr	Gln	Glu	Asn	
	860					865					870					
GGC	AAT	CAC	TCG	CGC	CAG	GAC	TCG	ATG	GAG	CCC	TGG	GAG	GAA	TAC	TCC	3772
Gly	Asn	His	Ser	Arg	Gln	Asp	Ser	Met	Glu	Pro	Trp	Glu	Glu	Tyr	Ser	
	875				880					885					890	
GAG	AAC	GGA	CCG	AAG	CCC	ACC	GTG	CTG	GAG	AGG	CGA	CGG	ATG	CGC	GAT	3820
Glu	Asn	Gly	Pro	Lys	Pro	Thr	Val	Leu	Glu	Arg	Arg	Arg	Met	Arg	Asp	
				895					900					905		
CAC	CAA	AGA	GTC	CTG	GCC	CAC	TCG	CTG	GTG	GGC	ACC	CCG	AAC	TAC	ATA	3868
His	Gln	Arg	Val	Leu	Ala	His	Ser	Leu	Val	Gly	Thr	Pro	Asn	Tyr	Ile	
			910					915					920			
GCT	CCC	GAG	GTG	CTG	GAG	AGG	AGT	GGG	TAC	ACG	CAG	CTG	TGC	GAC	TAC	3916
Ala	Pro	Glu	Val	Leu	Glu	Arg	Ser	Gly	Tyr	Thr	Gln	Leu	Cys	Asp	Tyr	
		925					930					935				

TGG AGC GTG GGC GTC ATC CTT TAC GAG ATG CTG GTG GGT CAG CCG CCC	3964
Trp Ser Val Gly Val Ile Leu Tyr Glu Met Leu Val Gly Gln Pro Pro	
940 945 950	
TTT CTG GCC AAC AGT CCG CTG GAA ACG CAA CAA AAG GTC ATC AAC TGG	4012
Phe Leu Ala Asn Ser Pr Leu lu Thr Gln Gln Lys Val Ile Asn Trp	
955 960 965 970	
GAG AAA ACG CTG CAT ATT CCG CCG CAG GCC GAG TTA TCC CGC GAG GCT	4060
Glu Lys Thr Leu His Ile Pro Pro Gln Ala Glu Leu Ser Arg Glu Ala	
975 980 985	
ACG GAC TTG ATA AGG AGG CTC TGT GCG TCG GCT GAC AAG CGG CTG GGC	4108
Thr Asp Leu Ile Arg Arg Leu Cys Ala Ser Ala Asp Lys Arg Leu Gly	
990 995 1000	
AAG AGC GTG GAC GAG GTC AAG AGC CAC GAC TTC TTC AAG GGC ATC GAC	4156
Lys Ser Val Asp Glu Val Lys Ser His Asp Phe Phe Lys Gly Ile Asp	
1005 1010 1015	
TTT GCG GAC ATG CGG AAG CAG AAA GCG CCC TAC ATA CCG GAA ATC AAG	4204
Phe Ala Asp Met Arg Lys Gln Lys Ala Pro Tyr Ile Pro Glu Ile Lys	
1020 1025 1030	
CAC CCA ACG GAC ACA TCC AAC TTT GAT CCC GTG GAT CCG GAG AAG CTG	4252
His Pro Thr Asp Thr Ser Asn Phe Asp Pro Val Asp Pro Glu Lys Leu	
1035 1040 1045 1050	
CGC TCG AAT GAC TCC ACC ATG AGC AGC GGC GAT GAT GTC GAC CAG AAT	4300
Arg Ser Asn Asp Ser Thr Met Ser Ser Gly Asp Asp Val Asp Gln Asn	
1055 1060 1065	
GAC CGC ACT TTC CAC GGC TTT TTC GAA TTT ACC TTC CGT CGC TTC TTC	4348
Asp Arg Thr Phe His Gly Phe Phe Glu Phe Thr Phe Arg Arg Phe Phe	
1070 1075 1080	
GAC GAC AAG CAG CCG CCG GAT ATG ACG GAC GAT CAG GCG CCG GTT TAC	4396
Asp Asp Lys Gln Pro Pro Asp Met Thr Asp Asp Gln Ala Pro Val Tyr	
1085 1090 1095	
GTC TGA AATGGATGCT CTCCATGTGC CCAACACCAA CACCCCGCCC CCGAATCATT	4452
Val *	
1100	
GTTAGTCAAA TAGTCACAAA AAGGGGATAG AAACCATTGA GTGGGCTTGC ATTGTAAAGG	4512
AAGCGTGGCT ATAGAATGAA ACTATCTATA TACATTATAT AAATTATAGG AGACAGTAGA	4572
GGCGGGAGCT ACGTATATAC ATACAAATAA TATACATATA TTTGATATAT ATATATATAT	4632
ATATGCCGTA GGGCATGAAC TGAATAAATA TAAAACGGAG CCGAGTAGAG ATGAAACGAG	4692
AGGAGCGAGT CAGGACCTTC GACCTTTAAC TGAACATAGT ATATCCTTGT GCACTACTAC	4752
TCCACAACAA ATATATATTT TTAAATTGTT AGAATTCAAA AGGGACCAAC TGGAAATCGA	4812
ACCTTTCTGG TGCTCAAAGC AAAGCAAAGC AAAGCAAAAC AAAACGCCTT AAATAAATG	4872
AGACGCGAAT TTACCCAACC ACTTCACTCC TCTCCTTTCT CCACCTCCGA TCGGTGGCCG	4932
GATTTCGAACT CAGCAGGCTG GTTGCATCCG GCCATCCCAT TGACTTCCCA TTCAGAATTG	4992
AGATTGCGAG GTGTGCGATG GAGAACGAAC GGAGACCAAA AGTCGCACGG CAGCGATATA	5052
AGCGGGTCTT ATAAGCCTAA TCTAAATCTA AACTGGGAGA ACAGGACCTA TGTATGTCCT	5112
GCTATCCAAT TCGTCTATCA CTGCTCTTCA TCTGTGTACG ACCCCACCC CCCCCCTCCC	5172

CATCCAAAAG AACAAACTTA GACGTAGCCT ATGTGAAAAG CTAGCAATGT TAGACCAACT	5232
TGTTGAATGC CAAATGAAAT TGTTTAGCCC CACGAGGAAA ACGCGGGGGA AATTCAACAC	5292
TTATTCTCTG ATAGCAAACG GAAAAGAAAG AAAGAAAAAA AAAACAGAA ACAGTACGAG	5352
AAAATTGTAA TCTTCTTAAT GTAATATTGT AAAGAACACG TTAATTGTAA TCTATGCTAG	5412
AGTTGTGTAG CGCCCTAAGA TGTTTTTTAG TTTATAGACC GCTAACCGTA ATCTAGTTTA	5472
ATTCCTAACA CTAAGCGAGA GTACAGTACA TTGGTTTTTT TGTGTGTCGT AGGTTTCGTTG	5532
GAAAATGCTT AACGGGAAAC GATTTGTTTT TCTCTTTAAT TAGCTTCAGT TTGTATGTGC	5592
GTGTGTTTTT ATTATGACTT ATATATAGTC CATCTGAATA TTCGTGGATG GAGCCTATTT	5652
TAAATGTGAG ATCGAGCTAA TTGAAGGAAA TACAAACAAA CTCTGTGTGC CTTGGCCAAT	5712
TAGTTTAC	5720

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1100 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	His	Pro	Ala	Gly	Glu	Lys	Arg	Gly	Gly	Arg	Pro	Asn	Asp	Lys	Tyr	1	5	10	15
Thr	Ala	Glu	Ala	Leu	Glu	Ser	Ile	Lys	Gln	Asp	Leu	Thr	Arg	Phe	Glu	20	25	30	
Val	Gln	Asn	Asn	His	Arg	Asn	Asn	Gln	Asn	Tyr	Thr	Pro	Leu	Arg	Tyr	35	40	45	
Thr	Ala	Thr	Asn	Gly	Arg	Asn	Asp	Ala	Leu	Thr	Pro	Asp	Tyr	His	His	50	55	60	
Ala	Lys	Gln	Pro	Met	Glu	Pro	Pro	Pro	Ser	Ala	Ser	Pro	Ala	Pro	Asp	65	70	75	80
Val	Val	Ile	Pro	Pro	Pro	Pro	Ala	Ile	Val	Gly	Gln	Pro	Gly	Ala	Gly	85	90	95	
Ser	Ile	Ser	Val	Ser	Gly	Val	Gly	Val	Gly	Val	Val	Gly	Val	Ala	Asn	100	105	110	
Gly	Arg	Val	Pro	Lys	Met	Met	Thr	Ala	Leu	Met	Pro	Asn	Lys	Leu	Ile	115	120	125	
Arg	Lys	Pro	Ser	Ile	Glu	Arg	Asp	Thr	Ala	Ser	Ser	His	Tyr	Leu	Arg	130	135	140	
Cys	Ser	Pro	Ala	Leu	Asp	Ser	Gly	Ala	Gly	Ser	Ser	Arg	Ser	Asp	Ser	145	150	155	160
Pro	His	Ser	His	His	Thr	His	Gln	Pro	Ser	Ser	Arg	Thr	Val	Gly	Asn	165	170	175	
Pro	Gly	Gly	Asn	Gly	Gly	Phe	Ser	Pro	Ser	Pro	Ser	Gly	Phe	Ser	Glu	180	185	190	

Val Ala Pro Pro Ala Pro Pro Pro Arg Asn Pro Thr Ala Ser Ser Ala  
 195 200 205  
 Ala Thr Pro Pro Pro Pro Val Pro Pro Thr Ser Gln Ala Tyr Val Lys  
 210 215 220  
 Arg Arg Ser Pro Ala Leu Asn Asn Arg Pro Pro Ala Ile Ala Pro Pro  
 225 230 235 240  
 Thr Gln Arg Gly Asn Ser Pro Val Ile Thr Gln Asn Gly Leu Lys Asn  
 245 250 255  
 Pro Gln Gln Gln Leu Thr Gln Gln Leu Lys Ser Leu Asn Leu Tyr Pro  
 260 265 270  
 Gly Gly Gly Ser Gly Ala Val Val Glu Pro Pro Pro Pro Tyr Leu Ile  
 275 280 285  
 Gln Gly Gly Ala Gly Gly Ala Ala Pro Pro Pro Pro Pro Pro Ser Tyr  
 290 295 300  
 Thr Ala Ser Met Gln Ser Arg Gln Ser Pro Thr Gln Ser Gln Gln Ser  
 305 310 315 320  
 Asp Tyr Arg Lys Ser Pro Ser Ser Gly Ile Tyr Ser Ala Thr Ser Ala  
 325 330 335  
 Gly Ser Pro Ser Pro Ile Thr Val Ser Leu Pro Pro Ala Pro Leu Ala  
 340 345 350  
 Lys Pro Gln Pro Arg Val Tyr Gln Ala Arg Ser Gln Gln Pro Ile Ile  
 355 360 365  
 Met Gln Ser Val Lys Ser Thr Gln Val Gln Lys Pro Val Leu Gln Thr  
 370 375 380  
 Ala Val Ala Arg Gln Ser Pro Ser Ser Ala Ser Ala Ser Asn Ser Pro  
 385 390 395 400  
 Val His Val Leu Ala Ala Pro Pro Ser Tyr Pro Gln Lys Ser Ala Ala  
 405 410 415  
 Val Val Gln Gln Gln Gln Gln Ala Ala Ala Ala Ala His Gln Gln Gln  
 420 425 430  
 His Gln His Gln Gln Ser Lys Pro Pro Thr Pro Thr Thr Pro Pro Leu  
 435 440 445  
 Val Gly Leu Asn Ser Lys Pro Asn Cys Leu Glu Pro Pro Ser Tyr Ala  
 450 455 460  
 Lys Ser Met Gln Ala Lys Ala Ala Thr Val Val Gln Gln Gln Gln Gln  
 465 470 475 480  
 Gln Gln Gln Gln Gln Gln Val Gln Gln Gln Gln Val Gln Gln Gln Gln  
 485 490 495  
 Gln Gln Gln Gln Gln Gln Leu Gln Ala Leu Arg Val Leu Gln Ala Gln  
 500 505 510  
 Ala Gln Arg Glu Arg Asp Gln Arg Glu Arg Glu Arg Asp Gln Gln Lys  
 515 520 525  
 Leu Ala Asn Gly Asn Pro Gly Arg Gln Met Leu Pro Pro Pro Pro Tyr  
 530 535 540  
 Gln Ser Asn Asn Asn Asn Asn Ser Glu Ile Lys Pro Pr Ser Cys Asn

545	550					555					560				
Asn Asn Asn Ile	Gln 565	Ile Ser Asn S r	Asn 570	Leu Ala Thr Thr	Pro 575	Pro									
Ile Pro Pro Ala	Lys 580	Tyr Asn Asn Asn	Ser 585	S r Asn Thr	Gly 590	Ala Asn									
Ser Ser Gly Gly	Ser 595	Asn Gly Ser	Thr 600	Gly Thr Thr	Ala 605	Ser Ser Ser									
Thr Ser Cys Lys	Lys 610	Ile Lys His	Ala 615	Ser Pro Ile	Pro 620	Glu Arg Lys									
Lys Ile Ser Lys	Glu 625	Lys Glu Glu	Glu 630	Arg Lys Glu	Phe 635	Arg Ile Arg									
Gln Tyr Ser Pro	Gln 645	Ala Phe Lys	Phe 650	Met Glu Gln	His 655	Ile Glu									
Asn Val Ile Lys	Ser 660	Tyr Arg Gln	Arg 665	Thr Tyr Arg	Lys 670	Asn Gln Leu									
Glu Lys Glu Met	His 675	Lys Val Gly	Leu 680	Pro Asp Gln	Thr 685	Gln Ile Glu									
Met Arg Lys Met	Leu 690	Asn Gln Lys	Glu 695	Ser Asn Tyr	Ile 700	Arg Leu Lys									
Arg Ala Lys Met	Asp 705	Lys Ser Met	Phe 710	Val Lys Leu	Lys 715	Pro Ile Gly									
Val Gly Ala Phe	Gly 725	Glu Val Thr	Leu 730	Val Ser Lys	Ile 735	Asp Thr Ser									
Asn His Leu Tyr	Ala 740	Met Lys Thr	Leu 745	Arg Lys Ala	Asp 750	Val Leu Lys									
Arg Asn Gln Val	Ala 755	His Val Lys	Ala 760	Glu Arg Asp	Ile 765	Leu Ala Glu									
Ala Asp Asn Asn	Trp 770	Val Val Lys	Leu 775	Tyr Tyr Ser	Phe 780	Gln Asp Lys									
Asp Asn Leu Tyr	Phe 785	Val Met Asp	Tyr 790	Ile Pro Gly	Gly 795	Asp Leu Met									
Ser Leu Leu Ile	Lys 805	Leu Gly Ile	Phe 810	Glu Glu Glu	Leu 815	Ala Arg Phe									
Tyr Ile Ala Glu	Val 820	Thr Cys Ala	Val 825	Asp Ser Val	His 830	Lys Met Gly									
Phe Ile His Arg	Asp 835	Ile Lys Pro	Asp 840	Asn Ile Leu	Ile 845	Asp Arg Asp									
Gly His Ile Lys	Leu 850	Thr Asp Phe	Gly 855	Leu Cys Thr	Gly 860	Phe Arg Trp									
Thr His Asn Ser	Lys 865	Tyr Tyr Gln	Glu 870	Asn Gly Asn	His 875	Ser Arg Gln									
Asp Ser Met Glu	Pro 885	Trp Glu Glu	Tyr 890	Ser Glu Asn	Gly 895	Pro Lys Pro									
Thr Val L u Glu	Arg 900	Arg Arg Arg	Met 905	Arg Asp His	Gln 910	Val Leu Ala									

His Ser Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu Val Leu Glu  
 915 920 925  
 Arg Ser Gly Tyr Thr Gln Leu Cys Asp Tyr Trp Ser Val Gly Val Ile  
 930 935 940  
 Leu Tyr Glu Met Leu Val Gly Gln Pro Pro Ph Leu Ala Asn Ser Pro  
 945 950 955 960  
 Leu Glu Thr Gln Gln Lys Val Ile Asn Trp Glu Lys Thr Leu His Ile  
 965 970 975  
 Pro Pro Gln Ala Glu Leu Ser Arg Glu Ala Thr Asp Leu Ile Arg Arg  
 980 985 990  
 Leu Cys Ala Ser Ala Asp Lys Arg Leu Gly Lys Ser Val Asp Glu Val  
 995 1000 1005  
 Lys Ser His Asp Phe Phe Lys Gly Ile Asp Phe Ala Asp Met Arg Lys  
 1010 1015 1020  
 Gln Lys Ala Pro Tyr Ile Pro Glu Ile Lys His Pro Thr Asp Thr Ser  
 1025 1030 1035 1040  
 Asn Phe Asp Pro Val Asp Pro Glu Lys Leu Arg Ser Asn Asp Ser Thr  
 1045 1050 1055  
 Met Ser Ser Gly Asp Asp Val Asp Gln Asn Asp Arg Thr Phe His Gly  
 1060 1065 1070  
 Phe Phe Glu Phe Thr Phe Arg Arg Phe Phe Asp Asp Lys Gln Pro Pro  
 1075 1080 1085  
 Asp Met Thr Asp Asp Gln Ala Pro Val Tyr Val \*  
 1090 1095 1100

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3984 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 231..3623

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACCTTTGGGT TGCTGGGACG GACTCTGGCC GCCTCAGCGT CCGCCCTCAG GCCCCGTGGCC 60  
 GCTGTCCAGG AGCTCTGCTC TCCCCTCCAG AGTTAATTAT TTATATTGTA AAGAATTTTA 120  
 ACAGTCCTGG GGACTTCCTT GAAGGATCAT TTCACTTTT GCTCAGAAGA AAGCTCTGGA 180  
 TCTATCAAAT AAAGAAGTCC TTCGTGTGGG CTACATATAT AGATGTTTTT ATG AAG 236  
 Met Lys  
 1  
 AGG AGT GAA AAG CCA GAA GGA TAT AGA CAA ATG AGG CCT AAG ACC TTT 284  
 Arg Ser Glu Lys Pro Glu Gly Tyr Arg Gln Met Arg Pro Lys Thr Phe  
 5 10 15

CCT	GCC	AGT	AAC	TAT	ACT	GTC	AGT	AGC	CGG	CAA	ATG	TTA	CAA	GAA	ATT	332
Pro	Ala	Ser	Asn	Tyr	Thr	Val	Ser	Ser	Arg	Gln	Met	Leu	Gln	Glu	Ile	
	20					25					30					
CGG	GAA	TCC	CTT	AGG	AAT	TTA	TCT	AAA	CCA	TCT	GAT	GCT	GCT	AAG	GCT	380
Arg	Glu	Ser	Leu	Arg	Asn	Leu	Ser	Lys	Pro	Ser	Asp	Ala	Ala	Lys	Ala	
	35				40					45					50	
GAG	CAT	AAC	ATG	AGT	AAA	ATG	TCA	ACC	GAA	GAT	CCT	CGA	CAA	GTC	AGA	428
Glu	His	Asn	Met	Ser	Lys	Met	Ser	Thr	Glu	Asp	Pro	Arg	Gln	Val	Arg	
				55					60					65		
AAT	CCA	CCC	AAA	TTT	GGG	ACG	CAT	CAT	AAA	GCC	TTG	CAG	GAA	ATT	CGA	476
Asn	Pro	Pro	Lys	Phe	Gly	Thr	His	His	Lys	Ala	Leu	Gln	Glu	Ile	Arg	
			70					75					80			
AAC	TCT	CTG	CTT	CCA	TTT	GCA	AAT	GAA	ACA	AAT	TCT	TCT	CGG	AGT	ACT	524
Asn	Ser	Leu	Leu	Pro	Phe	Ala	Asn	Glu	Thr	Asn	Ser	Ser	Arg	Ser	Thr	
		85					90					95				
TCA	GAA	GTT	AAT	CCA	CAA	ATG	CTT	CAA	GAC	TTG	CAA	GCT	GCT	GGA	TTT	572
Ser	Glu	Val	Asn	Pro	Gln	Met	Leu	Gln	Asp	Leu	Gln	Ala	Ala	Gly	Phe	
	100					105					110					
GAT	GAG	GAT	ATG	GTT	ATA	CAA	GCT	CTT	CAG	AAA	ACT	AAC	AAC	AGA	AGT	620
Asp	Glu	Asp	Met	Val	Ile	Gln	Ala	Leu	Gln	Lys	Thr	Asn	Asn	Arg	Ser	
	115				120					125					130	
ATA	GAA	GCA	GCA	ATT	GAA	TTC	ATT	AGT	AAA	ATG	AGT	TAC	CAA	GAT	CCT	668
Ile	Glu	Ala	Ala	Ile	Glu	Phe	Ile	Ser	Lys	Met	Ser	Tyr	Gln	Asp	Pro	
				135					140					145		
CGA	CGA	GAG	CAG	ATG	GCT	GCA	GCA	GCT	GCC	AGA	CCT	ATT	AAT	GCC	AGC	716
Arg	Arg	Glu	Gln	Met	Ala	Ala	Ala	Ala	Ala	Arg	Pro	Ile	Asn	Ala	Ser	
			150					155					160			
ATG	AAA	CCA	GGG	AAT	GTG	CAG	CAA	TCA	GTT	AAC	CGC	AAA	CAG	AGC	TGG	764
Met	Lys	Pro	Gly	Asn	Val	Gln	Gln	Ser	Val	Asn	Arg	Lys	Gln	Ser	Trp	
		165					170					175				
AAA	GGT	TCT	AAA	GAA	TCC	TTA	GTT	CCT	CAG	AGG	CAT	GGC	CCG	CCA	CTA	812
Lys	Gly	Ser	Lys	Glu	Ser	Leu	Val	Pro	Gln	Arg	His	Gly	Pro	Pro	Leu	
	180					185					190					
GGA	GAA	AGT	GTG	GCC	TAT	CAT	TCT	GAG	AGT	CCC	AAC	TCA	CAG	ACA	GAT	860
Gly	Glu	Ser	Val	Ala	Tyr	His	Ser	Glu	Ser	Pro	Asn	Ser	Gln	Thr	Asp	
	195				200					205					210	
GTA	GGA	AGA	CCT	TTG	TCT	GGA	TCT	GGT	ATA	TCA	GCA	TTT	GTT	CAA	GCT	908
Val	Gly	Arg	Pro	Leu	Ser	Gly	Ser	Gly	Ile	Ser	Ala	Phe	Val	Gln	Ala	
			215					220						225		
CAC	CCT	AGC	AAC	GGA	CAG	AGA	GTG	AAC	CCC	CCA	CCA	CCA	CCT	CAA	GTA	956
His	Pro	Ser	Asn	Gly	Gln	Arg	Val	Asn	Pro	Pro	Pro	Pro	Pro	Gln	Val	
			230					235					240			
AGG	AGT	GTT	ACT	CCT	CCA	CCA	CCT	CCA	AGA	GGC	CAG	ACT	CCC	CCT	CCA	1004
Arg	Ser	Val	Thr	Pro	Pro	Pro	Pro	Pro	Arg	Gly	Gln	Thr	Pro	Pro	Pro	
		245					250					255				
AGA	GGT	ACA	ACT	CCA	CCT	CCC	CCT	TCA	TGG	GAA	CCA	AAC	TCT	CAA	ACA	1052
Arg	Gly	Thr	Thr	Pro	Pro	Pro	Pro	Ser	Trp	Glu	Pro	Asn	Ser	Gln	Thr	
	260					265					270					
AAG	CGC	TAT	TCT	GGA	AAC	ATG	GAA	TAC	GTA	ATC	TCC	CGA	ATC	TCT	CCT	1100
Lys	Arg	Tyr	Ser	Gly	Asn	Met	Glu	Tyr	Val	Ile	Ser	Arg	Ile	Ser	Pro	
	275				280					285					290	



GTC	CCA	CCT	GGG	GCA	TGG	CAA	GAG	GGC	TAT	CCT	CCA	CCA	CCT	CTC	AAC	1148
Val	Pro	Pro	Gly	Ala	Trp	Gln	Glu	Gly	Tyr	Pr	Pro	Pro	Pro	Leu	Asn	
			295						300					305		
ACT	TCC	CCC	ATG	AAT	CCT	CCT	AAT	CAA	GGA	CAG	AGA	GGC	ATT	AGT	TCT	1196
Thr	Ser	Pro	M t	Asn	Pro	Pro	Asn	Gln	Gly	Gln	Arg	Gly	Ile	Ser	Ser	
			310					315					320			
GTT	CCT	GTT	GGC	AGA	CAA	CCA	ATC	ATC	ATG	CAG	AGT	TCT	AGC	AAA	TTT	1244
Val	Pro	Val	Gly	Arg	Gln	Pro	Ile	Ile	Met	Gln	Ser	Ser	Ser	Lys	Phe	
			325				330					335				
AAC	TTT	CCA	TCA	GGG	AGA	CCT	GGA	ATG	CAG	AAT	GGT	ACT	GGA	CAA	ACT	1292
Asn	Phe	Pro	Ser	Gly	Arg	Pro	Gly	Met	Gln	Asn	Gly	Thr	Gly	Gln	Thr	
	340					345					350					
GAT	TTC	ATG	ATA	CAC	CAA	AAT	GTT	GTC	CCT	GCT	GGC	ACT	GTG	AAT	CGG	1340
Asp	Phe	Met	Ile	His	Gln	Asn	Val	Val	Pro	Ala	Gly	Thr	Val	Asn	Arg	
	355				360					365					370	
CAG	CCA	CCA	CCT	CCA	TAT	CCT	CTG	ACA	GCA	GCT	AAT	GGA	CAA	AGC	CCT	1388
Gln	Pro	Pro	Pro	Pro	Tyr	Pro	Leu	Thr	Ala	Ala	Asn	Gly	Gln	Ser	Pro	
					375				380					385		
TCT	GCT	TTA	CAA	ACA	GGG	GGA	TCT	GCT	GCT	CCT	TCG	TCA	TAT	ACA	AAT	1436
Ser	Ala	Leu	Gln	Thr	Gly	Gly	Ser	Ala	Ala	Pro	Ser	Ser	Tyr	Thr	Asn	
			390					395					400			
GGA	AGT	ATT	CCT	CAG	TCT	ATG	ATG	GTG	CCA	AAC	AGA	AAT	AGT	CAT	AAC	1484
Gly	Ser	Ile	Pro	Gln	Ser	Met	Met	Val	Pro	Asn	Arg	Asn	Ser	His	Asn	
		405					410					415				
ATG	GAA	CTA	TAT	AAC	ATT	AGT	GTA	CCT	GGA	CTG	CAA	ACA	AAT	TGG	CCT	1532
Met	Glu	Leu	Tyr	Asn	Ile	Ser	Val	Pro	Gly	Leu	Gln	Thr	Asn	Trp	Pro	
	420					425					430					
CAG	TCA	TCT	TCT	GCT	CCA	GCC	CAG	TCA	TCC	CCG	AGC	AGT	GGG	CAT	GAA	1580
Gln	Ser	Ser	Ser	Ala	Pro	Ala	Gln	Ser	Ser	Pro	Ser	Ser	Gly	His	Glu	
	435				440					445					450	
ATC	CCT	ACA	TGG	CAA	CCT	AAC	ATA	CCA	GTG	AGG	TCA	AAT	TCT	TTT	AAT	1628
Ile	Pro	Thr	Trp	Gln	Pro	Asn	Ile	Pro	Val	Arg	Ser	Asn	Ser	Phe	Asn	
				455					460					465		
AAC	CCA	TTA	GGA	AAT	AGA	GCA	AGT	CAC	TCT	GCT	AAT	TCT	CAG	CCT	TCT	1676
Asn	Pro	Leu	Gly	Asn	Arg	Ala	Ser	His	Ser	Ala	Asn	Ser	Gln	Pro	Ser	
			470					475					480			
GCT	ACA	ACA	GTC	ACT	GCA	ATT	ACA	CCA	GCT	CCT	ATT	CAA	CAG	CCT	GTG	1724
Ala	Thr	Thr	Val	Thr	Ala	Ile	Thr	Pro	Ala	Pro	Ile	Gln	Gln	Pro	Val	
			485				490					495				
AAA	AGT	ATG	CGT	GTA	TTA	AAA	CCA	GAG	CTA	CAG	ACT	GCT	TTA	GCA	CCT	1772
Lys	Ser	Met	Arg	Val	Leu	Lys	Pro	Glu	Leu	Gln	Thr	Ala	Leu	Ala	Pro	
	500					505					510					
ACA	CAC	CCT	TCT	TGG	ATA	CCA	CAG	CCA	ATT	CAA	ACT	GTT	CAA	CCC	AGT	1820
Thr	His	Pro	Ser	Trp	Ile	Pro	Gln	Pro	Ile	Gln	Thr	Val	Gln	Pro	Ser	
	515				520					525				530		
CCT	TTT	CCT	GAG	GGA	ACC	GCT	TCA	AAT	GTG	ACT	GTG	ATG	CCA	CCT	GTT	1868
Pro	Phe	Pro	Glu	Gly	Thr	Ala	Ser	Asn	Val	Thr	Val	Met	Pro	Pro	Val	
				535					540					545		
GCT	GAA	GCT	CCA	AAC	TAT	CAA	GGA	CCA	CCA	CCA	CCC	TAC	CCA	AAA	CAT	1916
Ala	Glu	Ala	Pro	Asn	Tyr	Gln	Gly	Pro	Pro	Pro	Pro	Tyr	Pro	Lys	His	
			550				555						560			

CTG	CTG	CAC	CAA	AAC	CCA	TCT	GTT	CCT	CCA	TAC	AG	TCA	ATC	AGT	AAG	1964
Leu	Leu	His	ln	Asn	Pro	Ser	Val	Pro	Pro	Tyr	Glu	Ser	Ile	Ser	Lys	
		565					570					575				
CCT	AGC	AAA	GAG	GAT	CAG	CCA	AGC	TTG	CCC	AAG	GAA	GAT	GAG	AGT	GAA	2012
Pro	Ser	Lys	Glu	Asp	Gln	Pro	Ser	Leu	Pro	Lys	lu	Asp	Glu	Ser	Glu	
	580					585					590					
AAG	AGT	TAT	GAA	AAT	GTT	GAT	AGT	GGG	GAT	AAA	GAA	AAG	AAA	CAG	ATT	2060
Lys	Ser	Tyr	Glu	Asn	Val	Asp	Ser	Gly	Asp	Lys	Glu	Lys	Lys	Gln	Ile	
595					600					605					610	
ACA	ACT	TCA	CCT	ATT	ACT	GTT	AGG	AAA	AAC	AAG	AAA	GAT	GAA	GAG	CGA	2108
Thr	Thr	Ser	Pro	Ile	Thr	Val	Arg	Lys	Asn	Lys	Lys	Asp	Glu	Glu	Arg	
				615					620					625		
AGG	GAA	TCT	CGT	ATT	CAA	AGT	TAT	TCT	CCT	CAA	GCA	TTT	AAA	TTC	TTT	2156
Arg	Glu	Ser	Arg	Ile	Gln	Ser	Tyr	Ser	Pro	Gln	Ala	Phe	Lys	Phe	Phe	
			630					635					640			
ATG	GAG	CAA	CAT	GTA	GAA	AAT	GTA	CTC	AAA	TCT	CAT	CAG	CAG	CGT	CTA	2204
Met	Glu	Gln	His	Val	Glu	Asn	Val	Leu	Lys	Ser	His	Gln	Gln	Arg	Leu	
		645					650					655				
CAT	CGT	AAA	AAA	CAA	TTA	GAG	AAT	GAA	ATG	ATG	CGG	GTT	GGA	TTA	TCT	2252
His	Arg	Lys	Lys	Gln	Leu	Glu	Asn	Glu	Met	Met	Arg	Val	Gly	Leu	Ser	
	660					665					670					
CAA	GAT	GCC	CAG	GAT	CAA	ATG	AGA	AAG	ATG	CTT	TGC	CAA	AAA	GAA	TCT	2300
Gln	Asp	Ala	Gln	Asp	Gln	Met	Arg	Lys	Met	Leu	Cys	Gln	Lys	Glu	Ser	
675					680					685					690	
AAT	TAC	ATC	CGT	CTT	AAA	AGG	GCT	AAA	ATG	GAC	AAG	TCT	ATG	TTT	GTG	2348
Asn	Tyr	Ile	Arg	Leu	Lys	Arg	Ala	Lys	Met	Asp	Lys	Ser	Met	Phe	Val	
				695					700					705		
AAG	ATA	AAG	ACA	CTA	GGA	ATA	GGA	GCA	TTT	GGT	GAA	GTC	TGT	CTA	GCA	2396
Lys	Ile	Lys	Thr	Leu	Gly	Ile	Gly	Ala	Phe	Gly	Glu	Val	Cys	Leu	Ala	
			710					715					720			
AGA	AAA	GTA	GAT	ACT	AAG	GCT	TTG	TAT	GCA	ACA	AAA	ACT	CTT	CGA	AAG	2444
Arg	Lys	Val	Asp	Thr	Lys	Ala	Leu	Tyr	Ala	Thr	Lys	Thr	Leu	Arg	Lys	
		725					730					735				
AAA	GAT	GTT	CTT	CTT	CGA	AAT	CAA	GTC	GCT	CAT	GTT	AAG	GCT	GAG	AGA	2492
Lys	Asp	Val	Leu	Leu	Arg	Asn	Gln	Val	Ala	His	Val	Lys	Ala	Glu	Arg	
	740					745					750					
GAT	ATC	CTG	GCT	GAA	GCT	GAC	AAT	GAA	TGG	GTA	GTT	CGT	CTA	TAT	TAT	2540
Asp	Ile	Leu	Ala	Glu	Ala	Asp	Asn	Glu	Trp	Val	Val	Arg	Leu	Tyr	Tyr	
755					760					765					770	
TCA	TTC	CAA	GAT	AAG	GAC	AAT	TTA	TAC	TTT	GTA	ATG	GAC	TAC	ATT	CCT	2588
Ser	Phe	Gln	Asp	Lys	Asp	Asn	Leu	Tyr	Phe	Val	Met	Asp	Tyr	Ile	Pro	
				775					780					785		
GGG	GGT	GAT	ATG	ATG	AGC	CTA	TTA	ATT	AGA	ATG	GGC	ATC	TTT	CCA	GAA	2636
Gly	Gly	Asp	Met	Met	Ser	Leu	Leu	Ile	Arg	Met	Gly	Ile	Phe	Pro	Glu	
			790					795					800			
AGT	CTG	GCA	CGA	TTC	TAC	ATA	GCA	GAA	CTT	ACC	TGT	GCA	GTT	GAA	AGT	2684
Ser	Leu	Ala	Arg	Phe	Tyr	Ile	Ala	Glu	Leu	Thr	Cys	Ala	Val	Glu	Ser	
		805					810					815				
GTT	CAT	AAA	ATG	GGT	TTT	ATT	CAT	AGA	GAT	ATT	AAA	CCT	GAT	AAT	ATT	2732
Val	His	Lys	Met	Gly	Ph	Il	His	Arg	Asp	Ile	Lys	Pro	Asp	Asn	Ile	
	820					825					830					

TTG Leu 835	ATT Ile	GAT Asp	CGT Arg	GAT Asp	GGT Gly 840	CAT His	ATT Ile	AAA Lys	TTG Leu	ACT Thr 845	GAC Asp	TTT Phe	GGC Gly	CTC Leu	TGC Cys 850	2780
ACT Thr	GGC Gly	TTC Phe	AGA Arg	TGG Trp 855	ACA Thr	CAC His	GAT Asp	TCT Ser	AAG Lys 860	TAC Tyr	TAT Tyr	CAG Gln	AGT Ser	GGT Gly 865	GAC Asp	2828
CAT His	CCA Pro	CGG Arg	CAA Gln 870	GAT Asp	AGC Ser	ATG Met	GAT Asp	TTC Phe 875	AGT Ser	AAT Asn	GAA Glu	TGG Trp	GGG Gly 880	GAT Asp	CCC Pro	2876
TCA Ser	AGC Ser	TGT Cys 885	CGA Arg	TGT Cys	GGA Gly	GAC Asp	AGA Arg 890	CTG Leu	AAG Lys	CCA Pro	TTA Leu	GAG Glu 895	CGG Arg	AGA Arg	GCT Ala	2924
GCA Ala 900	CGC Arg	CAG Gln	CAC His	CAG Gln	CGA Arg	TGT Cys 905	CTA Leu	GCA Ala	CAT His	TCT Ser	TTG Leu 910	GTT Val	GGG Gly	ACT Thr	CCC Pro	2972
AAT Asn 915	TAT Tyr	ATT Ile	GCA Ala	CCT Pro	GAA Glu 920	GTG Val	TTG Leu	CTA Leu	CGA Arg	ACA Thr 925	GGA Gly	TAC Tyr	ACA Thr	CAG Gln	TTG Leu 930	3020
TGT Cys	GAT Asp	TGG Trp	TGG Trp	AGT Ser 935	GTT Val	GGT Gly	GTT Val	ATT Ile	CTT Leu 940	TTT Phe	GAA Glu	ATG Met	TTG Leu	GTG Val 945	GGA Gly	3068
CAA Gln	CCT Pro	CCT Pro	TTC Phe 950	TTG Leu	GCA Ala	CAA Gln	ACA Thr	CCA Pro 955	TTA Leu	GAA Glu	ACA Thr	CAA Gln	ATG Met 960	AAG Lys	GTT Val	3116
ATC Ile	AAC Asn 965	TGG Trp	CAA Gln	ACA Thr	TCT Ser	CTT Leu	CAC His 970	ATT Ile	CCA Pro	CCA Pro	CAA Gln	GCT Ala 975	AAA Lys	CTC Leu	AGT Ser	3164
CCT Pro 980	GAA Glu	GCT Ala	TCT Ser	GAT Asp	CTT Leu	ATT Ile 985	ATT Ile	AAA Lys	CTT Leu	TGC Cys	CGA Arg 990	GGA Gly	CCC Pro	GAA Glu	GAT Asp	3212
CGC Arg 995	TTA Leu	GGC Gly	AAG Lys	AAT Asn	GGT Gly 1000	GCT Ala	GAT Asp	GAA Glu	ATA Ile	AAA Lys 1005	GCT Ala	CAT His	CCA Pro	TTT Phe 1010	TTT Phe	3260
AAA Lys	ACA Thr	ATT Ile	GAC Asp 1015	TTC Phe	TCC Ser	AGT Ser	GAC Asp	CTG Leu	AGA Arg 1020	CAG Gln	CAG Gln	TCT Ser	GCT Ala	TCA Ser 1025	TAC Tyr	3308
ATT Ile	CCT Pro	AAA Lys	ATC Ile 1030	ACA Thr	CAC His	CCA Pro	ACA Thr	GAT Asp 1035	ACA Thr	TCA Ser	AAT Asn	TTT Phe	GAT Asp 1040	CCT Pro	GTT Val	3356
GAT Asp	CCT Pro	GAT Asp 1045	AAA Lys	TTA Leu	TGG Trp	AGT Ser	GAT Asp 1050	GAT Asp	AAC Asn	GAG Glu	GAA Glu	GAA Glu 1055	AAT Asn	GTA Val	AAT Asn	3404
GAC Asp 1060	ACT Thr	CTC Leu	AAT Asn	GGA Gly	TGG Trp	TAT Tyr	AAA Lys 1065	AAT Asn	GGA Gly	AAG Lys	CAT His 1070	CCT Pro	GAA Glu	CAT His	GCA Ala	3452
TTC Phe 1075	TAT Tyr	GAA Glu	TTT Phe	ACC Thr	TTC Phe	CGA Arg	AGG Arg	TTT Phe	TTT Phe	GAT Asp 1085	GAC Asp	AAT Asn	GGC Gly	TAC Tyr	CCA Pro 1090	3500
TAT Tyr	AAT Asn	TAT Tyr	CCG Pro	AAG Lys 1095	CCT Pro	ATT Ile	GAA Glu	TAT Tyr	GAA Glu 1100	TAC Tyr	ATT Ile	AAT Asn	TCA Ser	CAA Gln	GGC Gly 1105	3548

TCA GAG CAG CAG TCG GAT GAA GAT GAT CAA AAC ACA GGC TCA GAG ATT 3596  
 Ser Glu Gln In Ser Asp Glu Asp Asp Gln Asn Thr Gly S r Glu Il  
                   1110                  1115                  1120

AAA AAT CGC GAT CTA GTA TAT GTT TAA CACACTAGTA AATAAATGTA 3643  
 Lys Asn Arg Asp Leu Val Tyr Val \*

1125                  1130

ATGAGGATTT GTAAAAGGGC CTGAAATGCG AGGTGTTTTG AGGTTCTGAG AGTAAAATTA 3703

TGCAAATATG ACAGAGCTAT ATATGTGTGC TCTGTGTACA ATATTTTATT TTCCTAAATT 3763

ATGGGAAATC CTTTTAAAAT GTTAATTTAT TCCAGCCGTT TAAATCAGTA TTTAGAAAAA 3823

AATTGTTATA AGGAAAGTAA ATTATGAACT GAATATTATA GTCAGTTCTT GGTACTTAAA 3883

GTACTIONAAA TAAGTAGTGC TTTGTTTTAAA AGGAGAAACC TGGTATCTAT TTGTATATAT 3943

GCTAAATAAT TTTAAAATAC AAGAGTTTTT GAAATTTTTT T 3984

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1131 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Lys Arg Ser Glu Lys Pro Glu Gly Tyr Arg Gln Met Arg Pro Lys  
 1                  5                  10                  15

Thr Phe Pro Ala Ser Asn Tyr Thr Val Ser Ser Arg Gln Met Leu Gln  
                   20                  25                  30

Glu Ile Arg Glu Ser Leu Arg Asn Leu Ser Lys Pro Ser Asp Ala Ala  
                   35                  40                  45

Lys Ala Glu His Asn Met Ser Lys Met Ser Thr Glu Asp Pro Arg Gln  
                   50                  55                  60

Val Arg Asn Pro Pro Lys Phe Gly Thr His His Lys Ala Leu Gln Glu  
                   65                  70                  75                  80

Ile Arg Asn Ser Leu Leu Pro Phe Ala Asn Glu Thr Asn Ser Ser Arg  
                   85                  90                  95

Ser Thr Ser Glu Val Asn Pro Gln Met Leu Gln Asp Leu Gln Ala Ala  
                   100                  105                  110

Gly Phe Asp Glu Asp Met Val Ile Gln Ala Leu Gln Lys Thr Asn Asn  
                   115                  120                  125

Arg Ser Ile Glu Ala Ala Ile Glu Phe Ile Ser Lys Met Ser Tyr Gln  
                   130                  135                  140

Asp Pro Arg Arg Glu Gln Met Ala Ala Ala Ala Arg Pro Ile Asn  
                   145                  150                  155                  160

Ala Ser Met Lys Pro Gly Asn Val Gln Gln Ser Val Asn Arg Lys Gln  
                   165                  170                  175

Ser Trp Lys Gly Ser Lys Glu Ser Leu Val Pro Gln Arg His Gly Pro  
                   180                  185                  190

Pro Leu Gly Glu Ser Val Ala Tyr His S r Glu Ser Pro Asn Ser Gln  
 195 200 205  
 Thr Asp Val Gly Arg Pro Leu Ser Gly Ser Gly Ile Ser Ala Phe Val  
 210 215 220  
 Gln Ala His Pro Ser Asn Gly Gln Arg Val Asn Pro Pro Pro Pro Pro  
 225 230 235 240  
 Gln Val Arg Ser Val Thr Pro Pro Pro Pro Pro Arg Gly Gln Thr Pro  
 245 250 255  
 Pro Pro Arg Gly Thr Thr Pro Pro Pro Pro Ser Trp Glu Pro Asn Ser  
 260 265 270  
 Gln Thr Lys Arg Tyr Ser Gly Asn Met Glu Tyr Val Ile Ser Arg Ile  
 275 280 285  
 Ser Pro Val Pro Pro Gly Ala Trp Gln Glu Gly Tyr Pro Pro Pro Pro  
 290 295 300  
 Leu Asn Thr Ser Pro Met Asn Pro Pro Asn Gln Gly Gln Arg Gly Ile  
 305 310 315 320  
 Ser Ser Val Pro Val Gly Arg Gln Pro Ile Ile Met Gln Ser Ser Ser  
 325 330 335  
 Lys Phe Asn Phe Pro Ser Gly Arg Pro Gly Met Gln Asn Gly Thr Gly  
 340 345 350  
 Gln Thr Asp Phe Met Ile His Gln Asn Val Val Pro Ala Gly Thr Val  
 355 360 365  
 Asn Arg Gln Pro Pro Pro Pro Tyr Pro Leu Thr Ala Ala Asn Gly Gln  
 370 375 380  
 Ser Pro Ser Ala Leu Gln Thr Gly Gly Ser Ala Ala Pro Ser Ser Tyr  
 385 390 395 400  
 Thr Asn Gly Ser Ile Pro Gln Ser Met Met Val Pro Asn Arg Asn Ser  
 405 410 415  
 His Asn Met Glu Leu Tyr Asn Ile Ser Val Pro Gly Leu Gln Thr Asn  
 420 425 430  
 Trp Pro Gln Ser Ser Ser Ala Pro Ala Gln Ser Ser Pro Ser Ser Gly  
 435 440 445  
 His Glu Ile Pro Thr Trp Gln Pro Asn Ile Pro Val Arg Ser Asn Ser  
 450 455 460  
 Phe Asn Asn Pro Leu Gly Asn Arg Ala Ser His Ser Ala Asn Ser Gln  
 465 470 475 480  
 Pro Ser Ala Thr Thr Val Thr Ala Ile Thr Pro Ala Pro Ile Gln Gln  
 485 490 495  
 Pro Val Lys Ser Met Arg Val Leu Lys Pro Glu Leu Gln Thr Ala Leu  
 500 505 510  
 Ala Pro Thr His Pro Ser Trp Ile Pro Gln Pro Ile Gln Thr Val Gln  
 515 520 525  
 Pro Ser Pro Phe Pro Glu Gly Thr Ala Ser Asn Val Thr Val Met Pro  
 530 535 540  
 Pro Val Ala Glu Ala Pr Asn Tyr Gln Gly Pro Pro Pro Pro Tyr Pro

545	550	555	560
Lys His Leu Leu	His Gln Asn Pro Ser	Val Pro Pro Tyr Glu	S r Ile
	565	570	575
Ser Lys Pro Ser	Lys Glu Asp Gln Pro	Ser Leu Pro Lys Glu	Asp Glu
	580	585	590
Ser Glu Lys Ser	Tyr Glu Asn Val Asp	Ser Gly Asp Lys Glu	Lys Lys
	595	600	605
Gln Ile Thr Thr	Ser Pro Ile Thr Val Arg	Lys Asn Lys Lys Asp	Glu
	610	615	620
Glu Arg Arg Glu	Ser Arg Ile Gln Ser Tyr	Ser Pro Gln Ala Phe	Lys
	625	630	635
Phe Phe Met Glu	Gln His Val Glu Asn	Val Leu Lys Ser His	Gln Gln
	645	650	655
Arg Leu His Arg	Lys Lys Gln Leu Glu	Asn Glu Met Met Arg	Val Gly
	660	665	670
Leu Ser Gln Asp	Ala Gln Asp Gln Met	Arg Lys Met Leu Cys	Gln Lys
	675	680	685
Glu Ser Asn Tyr	Ile Arg Leu Lys Arg	Ala Lys Met Asp	Lys Ser Met
	690	695	700
Phe Val Lys Ile	Lys Thr Leu Gly Ile	Gly Ala Phe Gly Glu	Val Cys
	705	710	715
Leu Ala Arg Lys	Val Asp Thr Lys Ala	Leu Tyr Ala Thr Lys	Thr Leu
	725	730	735
Arg Lys Lys Asp	Val Leu Leu Arg Asn	Gln Val Ala His Val	Lys Ala
	740	745	750
Glu Arg Asp Ile	Leu Ala Glu Ala Asp	Asn Glu Trp Val Val	Arg Leu
	755	760	765
Tyr Tyr Ser Phe	Gln Asp Lys Asp Asn	Leu Tyr Phe Val Met	Asp Tyr
	770	775	780
Ile Pro Gly Gly	Asp Met Met Ser Leu	Leu Ile Arg Met Gly	Ile Phe
	785	790	795
Pro Glu Ser Leu	Ala Arg Phe Tyr Ile	Ala Glu Leu Thr Cys	Ala Val
	805	810	815
Glu Ser Val His	Lys Met Gly Phe Ile	His Arg Asp Ile Lys	Pro Asp
	820	825	830
Asn Ile Leu Ile	Asp Arg Asp Gly His	Ile Lys Leu Thr Asp	Phe Gly
	835	840	845
Leu Cys Thr Gly	Phe Arg Trp Thr His	Asp Ser Lys Tyr Tyr	Gln Ser
	850	855	860
Gly Asp His Pro	Arg Gln Asp Ser Met	Asp Phe Ser Asn Glu	Trp Gly
	865	870	875
Asp Pro Ser Ser	Cys Arg Cys Gly Asp	Arg Leu Lys Pro Leu	Glu Arg
	885	890	895
Arg Ala Ala Arg	Gln His Gln Arg Cys	Leu Ala His Ser Leu	Val Gly
	900	905	910

Thr Pro Asn Tyr Il Ala Pro Glu Val Leu Leu Arg Thr ly Tyr Thr  
 915 920 925  
 Gln Leu Cys Asp Trp Trp Ser Val Gly Val Ile Leu Phe Glu Met Leu  
 930 935 940  
 Val Gly Gln Pro Pro Phe Leu Ala Gln Thr Pro L u Glu Thr Gln Met  
 945 950 955 960  
 Lys Val Ile Asn Trp Gln Thr Ser Leu His Ile Pro Pro Gln Ala Lys  
 965 970 975  
 Leu Ser Pro Glu Ala Ser Asp Leu Ile Ile Lys Leu Cys Arg Gly Pro  
 980 985 990  
 Glu Asp Arg Leu Gly Lys Asn Gly Ala Asp Glu Ile Lys Ala His Pro  
 995 1000 1005  
 Phe Phe Lys Thr Ile Asp Phe Ser Ser Asp Leu Arg Gln Gln Ser Ala  
 1010 1015 1020  
 Ser Tyr Ile Pro Lys Ile Thr His Pro Thr Asp Thr Ser Asn Phe Asp  
 1025 1030 1035 1040  
 Pro Val Asp Pro Asp Lys Leu Trp Ser Asp Asp Asn Glu Glu Glu Asn  
 1045 1050 1055  
 Val Asn Asp Thr Leu Asn Gly Trp Tyr Lys Asn Gly Lys His Pro Glu  
 1060 1065 1070  
 His Ala Phe Tyr Glu Phe Thr Phe Arg Arg Phe Phe Asp Asp Asn Gly  
 1075 1080 1085  
 Tyr Pro Tyr Asn Tyr Pro Lys Pro Ile Glu Tyr Glu Tyr Ile Asn Ser  
 1090 1095 1100  
 Gln Gly Ser Glu Gln Gln Ser Asp Glu Asp Asp Gln Asn Thr Gly Ser  
 1105 1110 1115 1120  
 Glu Ile Lys Asn Arg Asp Leu Val Tyr Val \*  
 1125 1130

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3213 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2889

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTG CAA CAT TCA ATT AAC CGA AAA CAA AGC TGG AAA GGT TCT AAA GAG	48
Val Gln His Ser Ile Asn Arg Lys Gln Ser Trp Lys Gly Ser Lys Glu	
1 5 10 15	
TCT CTA GTT CCT CAG AGA CAC GGC CCA TCT CTA GGA GAA AAT GTG GTT	96
Ser Leu Val Pro Gln Arg His Gly Pro Ser Leu Gly Glu Asn Val Val	
20 25 30	

TAT	CGT	TCT	GAA	AGC	CCC	AAC	TCA	CAG	GCG	GAT	GTA	GGA	AGA	CCT	CTG	144
Tyr	Arg	Ser	Glu	Ser	Pro	Asn	Ser	Gln	Ala	Asp	Val	Gly	Arg	Pro	Leu	
		35					40					45				
TCT	GGA	TCC	GGC	ATT	GCA	GCA	TTT	GCT	CAA	GCT	CAC	CCA	AGC	AAT	GGA	192
Ser	Gly	Ser	Gly	Ile	Ala	Ala	Ph	Ala	Gln	Ala	His	Pr	Ser	Asn	Gly	
	50					55					60					
CAG	AGA	GTG	AAC	CCC	CCA	CCA	CCA	CCT	CAA	GTT	AGG	AGT	GTT	ACT	CCT	240
Gln	Arg	Val	Asn	Pro	Pro	Pro	Pro	Pro	Gln	Val	Arg	Ser	Val	Thr	Pro	
	65				70					75					80	
CCA	CCA	CCT	CCG	AGA	GGC	CAG	ACC	CCA	CCT	CCC	CGA	GGC	ACC	ACT	CCC	288
Pro	Pro	Pro	Pro	Arg	Gly	Gln	Thr	Pro	Pro	Pro	Arg	Gly	Thr	Thr	Pro	
				85					90					95		
CCT	CCC	CCC	TCA	TGG	GAA	CCA	AGC	TCT	CAG	ACA	AAG	CGC	TAC	TCT	GGG	336
Pro	Pro	Pro	Ser	Trp	Glu	Pro	Ser	Ser	Gln	Thr	Lys	Arg	Tyr	Ser	Gly	
			100					105					110			
AAC	ATG	GAG	TAC	GTA	ATC	TCC	CGA	ATC	TCC	CCT	GTT	CCA	CCT	GGG	GCG	384
Asn	Met	Glu	Tyr	Val	Ile	Ser	Arg	Ile	Ser	Pro	Val	Pro	Pro	Gly	Ala	
		115					120					125				
TGG	CAG	GAG	GGG	TAC	CCT	CCA	CCA	CCT	CTT	ACC	ACT	TCT	CCC	ATG	AAT	432
Trp	Gln	Glu	Gly	Tyr	Pro	Pro	Pro	Pro	Leu	Thr	Thr	Ser	Pro	Met	Asn	
	130					135					140					
CCC	CCT	AGC	CAG	GCT	CAG	AGG	GCC	ATT	AGT	TCT	GTT	CCA	GTT	GGT	AGA	480
Pro	Pro	Ser	Gln	Ala	Gln	Arg	Ala	Ile	Ser	Ser	Val	Pro	Val	Gly	Arg	
	145				150					155					160	
CAA	CCC	ATC	ATC	ATG	CAG	AGT	ACT	AGC	AAA	TTT	AAC	TTT	ACA	CCA	GGG	528
Gln	Pro	Ile	Ile	Met	Gln	Ser	Thr	Ser	Lys	Phe	Asn	Phe	Thr	Pro	Gly	
				165					170					175		
CGA	CCT	GGA	GTT	CAG	AAT	GGT	GGT	GGT	CAG	TCT	GAT	TTT	ATC	GTG	CAC	576
Arg	Pro	Gly	Val	Gln	Asn	Gly	Gly	Gly	Gln	Ser	Asp	Phe	Ile	Val	His	
			180					185					190			
CAA	AAT	GTC	CCC	ACT	GGT	TCT	GTG	ACT	CGG	CAG	CCA	CCA	CCT	CCA	TAT	624
Gln	Asn	Val	Pro	Thr	Gly	Ser	Val	Thr	Arg	Gln	Pro	Pro	Pro	Pro	Tyr	
		195					200				205					
CCT	CTG	ACC	CCA	GCT	AAT	GGA	CAA	AGC	CCC	TCT	GCT	TTA	CAA	ACA	GGG	672
Pro	Leu	Thr	Pro	Ala	Asn	Gly	Gln	Ser	Pro	Ser	Ala	Leu	Gln	Thr	Gly	
	210					215					220					
GCT	TCT	GCT	GCT	CCA	CCA	TCA	TTC	GCC	AAT	GGA	AAC	GTT	CCT	CAG	TCG	720
Ala	Ser	Ala	Ala	Pro	Pro	Ser	Phe	Ala	Asn	Gly	Asn	Val	Pro	Gln	Ser	
	225				230					235					240	
ATG	ATG	GTG	CCC	AAC	AGG	AAC	AGT	CAT	AAC	ATG	GAG	CTT	TAT	AAT	ATT	768
Met	Met	Val	Pro	Asn	Arg	Asn	Ser	His	Asn	Met	Glu	Leu	Tyr	Asn	Ile	
				245					250					255		
AAT	GTC	CCT	GGA	CTG	CAA	ACA	GCC	TGG	CCC	CAG	TCG	TCT	TCT	GCT	CCT	816
Asn	Val	Pro	Gly	Leu	Gln	Thr	Ala	Trp	Pro	Gln	Ser	Ser	Ser	Ala	Pro	
			260					265					270			
GCG	CAG	TCA	TCC	CCA	AGC	GGT	GGG	CAT	GAA	ATT	CCT	ACA	TGG	CAA	CCT	864
Ala	Gln	Ser	Ser	Pro	Ser	Gly	Gly	His	Glu	Ile	Pro	Thr	Trp	Gln	Pro	
		275				280						285				
AAC	ATA	CCA	GTG	AGG	TCA	AAT	TCT	TTT	AAT	AAC	CCA	TTA	GGA	AGT	AGA	912
Asn	Ile	Pro	Val	Arg	Ser	Asn	Ser	Phe	Asn	Asn	Pro	Leu	Gly	Ser	Arg	
	290					295					300					



GCA Ala 305	AGT Ser	CAC His	TCT Ser	GCT Ala	AAT Asn 310	TCT Ser	CAG Gln	CCT Pro	TCT S r	GCC Ala 315	ACT Thr	ACA Thr	GTC Val	ACT Thr	GCC Ala 320	960
ATC Ile	ACA Thr	CCC Pro	GCT Ala	CCT Pro 325	ATT Ile	CAA Gln	CAG Gln	CCC Pro	GTG Val 330	AAA Lys	AGC Ser	ATG Met	CGC Arg	GTC Val 335	CTG Leu	1008
AAA Lys	CCA Pro	GAG Glu	CTG Leu 340	CAG Gln	ACT Thr	GCT Ala	TTA Leu 345	GCC Ala 345	CCA Pro	ACC Thr	CAT His	CCT Pro	TCT Ser 350	TGG Trp	ATG Met	1056
CCA Pro	CAG Gln	CCA Pro	GTT Val 355	CAG Gln	ACT Thr	GTT Val	CAG Gln 360	CCT Pro	ACC Thr	CCT Pro	TTT Phe	TCT Ser 365	GAG Glu	GGT Gly	ACA Thr	1104
GCT Ala 370	TCA Ser	AGT Ser	GTG Val	CCT Pro	GTC Val	ATC Ile 375	CCA Pro	CCT Pro	GTT Val	GCT Ala	GAA Glu 380	GCT Ala	CCA Pro	AGC Ser	TAT Tyr	1152
CAA Gln 385	GGT Gly	CCA Pro	CCA Pro	CCG Pro 390	CCT Pro	TAT Tyr	CCA Pro	AAA Lys	CAT His	CTG Leu 395	CTA Leu	CAC His	CAA Gln	AAC Asn	CCA Pro 400	1200
TCT Ser	GTC Val	CCT Pro	CCA Pro	TAT Tyr 405	GAG Glu	TCA Ser	GTA Val	AGT Ser	AAG Lys 410	CCC Pro	TGC Cys	AAA Lys	GAT Asp	GAA Glu 415	CAG Gln	1248
CCT Pro	AGC Ser	TTA Leu	CCC Pro 420	AAG Lys	GAA Glu	GAT Asp	GAT Asp	AGT Ser 425	GAG Glu	AAG Lys	AGT Ser	GCG Ala	GAC Asp 430	AGT Ser	GGT Gly	1296
GAC Asp	TCT Ser	GGG Gly 435	GAT Asp	AAA Lys	GAA Glu	AAG Lys	AAA Lys 440	CAG Gln	ATT Ile	ACA Thr	ACT Thr	TCA Ser 445	CCT Pro	ATC Ile	ACT Thr	1344
GTT Val 450	CGG Arg	AAA Lys	AAC Asn	AAG Lys	AAA Lys	GAT Asp 455	GAA Glu	GAA Glu	CGA Arg	AGA Arg	GAG Glu 460	TCT Ser	CGG Arg	ATT Ile	CAG Gln	1392
AGT Ser 465	TAC Tyr	TCC Ser	CCA Pro	CAG Gln	GCC Ala 470	TTT Phe	AAG Lys	TTC Phe	TTC Phe	ATG Met 475	GAG Glu	CAG Gln	CAC His	GTA Val	GAG Glu 480	1440
AAC Asn	GTC Val	CTG Leu	AAG Lys 485	TCT Ser	CAT His	CAG Gln	CAG Gln	CGT Arg 490	CTG Leu 490	CAT His	CGG Arg	AAG Lys	AAG Lys	CAG Gln 495	CTA Leu	1488
GAA Glu	AAT Asn	GAA Glu	ATG Met 500	ATG Met	CGG Arg	GTT Val	GGA Gly 505	TTA Leu 505	TCT Ser	CAA Gln	GAT Asp	GCC Ala 510	CAG Gln 510	GAT Asp	CAA Gln	1536
ATG Met	AGA Arg	AAG Lys 515	ATG Met	CTT Leu	TGC Cys	CAG Gln	AAA Lys 520	GAG Glu	TCT Ser	AAC Asn	TAT Tyr	ATT Ile 525	CGT Arg	CTT Leu	AAA Lys	1584
AGG Arg 530	GCT Ala	AAA Lys	ATG Met	GAC Asp	AAG Lys	TCT Ser 535	ATG Met	TTT Phe	GTA Val	AAG Lys 540	ATA Ile	AAG Lys	ACA Thr	TTA Leu	GGA Gly	1632
ATA Ile 545	GGA Gly	GCG Ala	TTT Phe	GGT Gly	GAA Glu 550	GTC Val	TGT Cys	CTA Leu	GCA Ala	AGA Arg 555	AAA Lys	GTC Val	GAT Asp	ACT Thr	AAA Lys 560	1680
GCT Ala	TTG L u	TAT Tyr	GCA Ala	ACA Thr 565	AAG Lys	ACT Thr	CTT Leu	CGA Arg	AAG Lys 570	AAA Lys	GAC Asp	GTT Val	CTG L u	CTC Leu	CGA Arg 575	1728

AAT CAG GTG GCT CAT GTG AAA GCG GAG AGG GAT ATC CTA GCA GAA GCC	1776
Asn Gln Val Ala His Val Lys Ala Glu Arg Asp Ile Leu Ala Glu Ala	
580 585 590	
GAC AAT GA TGG TG GTC CGC CTG TAC TAC TCT TTC CAG GAC AAG GAC	1824
Asp Asn lu Trp Val Val Arg Leu Tyr Tyr Ser Ph Gln Asp Lys Asp	
595 600 605	
AAC TTG TAC TTT GTG ATG GAC TAC ATT CCT GGG GGG GAT ATG ATG AGC	1872
Asn Leu Tyr Phe Val Met Asp Tyr Ile Pro Gly Gly Asp Met Met Ser	
610 615 620	
CTA TTA ATT AGA ATG GGC ATC TTT CCT GAA AAT CTG GCA CGA TTC TAC	1920
Leu Leu Ile Arg Met Gly Ile Phe Pro Glu Asn Leu Ala Arg Phe Tyr	
625 630 635 640	
ATA GCA GAA CTT ACC TGT GCA GTT GAA AGT GTT CAT AAA ATG GGT TTT	1968
Ile Ala Glu Leu Thr Cys Ala Val Glu Ser Val His Lys Met Gly Phe	
645 650 655	
ATT CAT AGA GAT ATT AAA CCT GAT AAC ATT TTG ATT GAC CGT GAT GGC	2016
Ile His Arg Asp Ile Lys Pro Asp Asn Ile Leu Ile Asp Arg Asp Gly	
660 665 670	
CAT ATT AAA TTG ACT GAC TTT GGC TTG TGC ACT GGC TTC AGA TGG ACA	2064
His Ile Lys Leu Thr Asp Phe Gly Leu Cys Thr Gly Phe Arg Trp Thr	
675 680 685	
CAT GAC TCC AAG TAC TAC CAG AGT GGG GAT CAC CCA CGG CAA GAT AGC	2112
His Asp Ser Lys Tyr Tyr Gln Ser Gly Asp His Pro Arg Gln Asp Ser	
690 695 700	
ATG GAT TTC AGT AAC GAA TGG GGA GAT CCT TCC AAT TGT CGG TGT GGG	2160
Met Asp Phe Ser Asn Glu Trp Gly Asp Pro Ser Asn Cys Arg Cys Gly	
705 710 715 720	
GAC AGA CTG AAG CCA CTG GAG CGG AGA GCT GCT CGC CAG CAC CAG CGA	2208
Asp Arg Leu Lys Pro Leu Glu Arg Arg Ala Ala Arg Gln His Gln Arg	
725 730 735	
TGT CTA GCC CAT TCT CTG GTT GGG ACT CCC AAT TAT ATT GCA CCT GAA	2256
Cys Leu Ala His Ser Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu	
740 745 750	
GTG CTA CTG CGA ACA GGA TAT ACA CAG CTG TGT GAC TGG TGG AGT GTT	2304
Val Leu Leu Arg Thr Gly Tyr Thr Gln Leu Cys Asp Trp Trp Ser Val	
755 760 765	
GGT GTT ATT CTT TGT GAA ATG TTG GTG GGA CAA CCT CCT TTC TTG GCA	2352
Gly Val Ile Leu Cys Glu Met Leu Val Gly Gln Pro Pro Phe Leu Ala	
770 775 780	
CAA ACC CCA TTA GAA ACA CAA ATG AAG GTT ATC ATC TGG CAA ACT TCT	2400
Gln Thr Pro Leu Glu Thr Gln Met Lys Val Ile Ile Trp Gln Thr Ser	
785 790 795 800	
CTA CAC ATC CCT CCT CAA GCT AAG CTG AGT CCT GAA GCC TCT GAC CTC	2448
Leu His Ile Pro Pro Gln Ala Lys Leu Ser Pro Glu Ala Ser Asp Leu	
805 810 815	
ATT ATC AAA CTG TGT CGA GGA CCA GAA GAC CGC CTC GGC AAG AAC GGT	2496
Ile Ile Lys Leu Cys Arg Gly Pro Glu Asp Arg Leu Gly Lys Asn Gly	
820 825 830	
GCT GAT GAG ATA AAG GCT CAT CCA TTT TTT AAG ACC ATC GAT TTC TCT	2544
Ala Asp Glu Ile Lys Ala His Pro Phe Phe Lys Thr Ile Asp Phe Ser	
835 840 845	

AGT GAT CTG AGA CAG CAG TCT GCT TCA TAC ATC CCT AAA ATC ACG CAT	2592
S r Asp Leu Arg Gln Gln Ser Ala Ser Tyr Ile Pro Lys Ile Thr His	
850 855 860	
CCA ACA GAT ACA TCC AAT TTC GAC CCT GTT GAT CCT GAT AAA TTG TGG	2640
Pro Thr Asp Thr Ser Asn Phe Asp Pro Val Asp Pro Asp Lys Leu Trp	
865 870 875 880	
AGC GAT GGC AGC GAG GAG GAA AAT ATC AGT GAC ACT CTG AGC GGA TGG	2688
Ser Asp Gly Ser Glu Glu Glu Asn Ile Ser Asp Thr Leu Ser Gly Trp	
885 890 895	
<del>TAT AAA AAT GGG AAG CAC CCC GAG CAC GCT TTC TAT GAG TTC ACC TTT</del>	2736
<del>Tyr Lys Asn Gly Lys His Pro Glu His Ala Phe Tyr Glu Phe Thr Phe</del>	
<del>900 905 910</del>	
CGG AGG TTT TTT GAT GAC AAT GGC TAC CCA TAT AAT TAT CCA AAG CCT	2784
Arg Arg Phe Phe Asp Asp Asn Gly Tyr Pro Tyr Asn Tyr Pro Lys Pro	
915 920 925	
ATT GAG TAT GAA TAC ATT CAT TCA CAG GGC TCA GAA CAA CAG TCT GAT	2832
Ile Glu Tyr Glu Tyr Ile His Ser Gln Gly Ser Glu Gln Gln Ser Asp	
930 935 940	
GAA GAT GAT CAA CAC ACA AGC TCC GAT GGA AAC AAC CGA GAT CTA GTG	2880
Glu Asp Asp Gln His Thr Ser Ser Asp Gly Asn Asn Arg Asp Leu Val	
945 950 955 960	
TAT GTT TAA TAACTAGGA GATCATTGTA AGAATTTGCA AGAGGCCTGA	2929
Tyr Val *	
AGTGCAGGGG TTTTGAAGT TTTGAGAAAA TTATGCAAAT GTGACAGAGT TTGTGTGCTC	2989
TGTGTACAAT ATTTTATTTT CCTAAGTTAT GGGAAATTGT TTTAAAATGT TAATTTATTC	3049
CACCCTTTTA ATTCAGTAAT TTAGAAAAAA TTGTTATAAG GAAAGTAAAT TATGAACTGA	3109
GTATTATAGT CAATTCTTGG TACTTAAAGT ACTTAAAAAG AGAAGCCTGG TATCTTTTGT	3169
ATATATAATA AATAATTTTA AAATCCCAA AAAAAAAAAA AAAA	3213

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 963 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Val Gln His Ser Ile Asn Arg Lys Gln Ser Trp Lys Gly Ser Lys Glu  
1 5 10 15  
Ser Leu Val Pro Gln Arg His Gly Pro Ser Leu Gly Glu Asn Val Val  
20 25 30  
Tyr Arg Ser Glu Ser Pro Asn Ser Gln Ala Asp Val Gly Arg Pro Leu  
35 40 45  
Ser Gly Ser Gly Ile Ala Ala Phe Ala Gln Ala His Pro Ser Asn Gly  
50 55 60  
Gln Arg Val Asn Pro Pro Pro Pro Gln Val Arg Ser Val Thr Pro

65						70						75						80
Pro	Pro	Pro	Pro	Arg	Gly	Gln	Thr	Pro	Pro	Pro	Arg	Gly	Thr	Thr	Pro			
				85					90					95				
Pro	Pro	Pro	S r	Trp	lu	Pro	Ser	S r	Gln	Thr	Lys	Arg	Tyr	Ser	ly			
			100					105					110					
Asn	Met	Glu	Tyr	Val	Ile	Ser	Arg	Ile	Ser	Pro	Val	Pro	Pro	Gly	Ala			
		115					120					125						
Trp	Gln	Glu	Gly	Tyr	Pro	Pro	Pro	Pro	Leu	Thr	Thr	Ser	Pro	Met	Asn			
	130					135					140							
Pro	Pro	Ser	Gln	Ala	Gln	Arg	Ala	Ile	Ser	Ser	Val	Pro	Val	Gly	Arg			
145					150					155					160			
Gln	Pro	Ile	Ile	Met	Gln	Ser	Thr	Ser	Lys	Phe	Asn	Phe	Thr	Pro	Gly			
				165					170					175				
Arg	Pro	Gly	Val	Gln	Asn	Gly	Gly	Gly	Gln	Ser	Asp	Phe	Ile	Val	His			
			180					185					190					
Gln	Asn	Val	Pro	Thr	Gly	Ser	Val	Thr	Arg	Gln	Pro	Pro	Pro	Pro	Tyr			
		195					200					205						
Pro	Leu	Thr	Pro	Ala	Asn	Gly	Gln	Ser	Pro	Ser	Ala	Leu	Gln	Thr	Gly			
	210					215					220							
Ala	Ser	Ala	Ala	Pro	Pro	Ser	Phe	Ala	Asn	Gly	Asn	Val	Pro	Gln	Ser			
225					230					235					240			
Met	Met	Val	Pro	Asn	Arg	Asn	Ser	His	Asn	Met	Glu	Leu	Tyr	Asn	Ile			
				245					250					255				
Asn	Val	Pro	Gly	Leu	Gln	Thr	Ala	Trp	Pro	Gln	Ser	Ser	Ser	Ala	Pro			
			260					265					270					
Ala	Gln	Ser	Ser	Pro	Ser	Gly	Gly	His	Glu	Ile	Pro	Thr	Trp	Gln	Pro			
		275					280					285						
Asn	Ile	Pro	Val	Arg	Ser	Asn	Ser	Phe	Asn	Asn	Pro	Leu	Gly	Ser	Arg			
	290					295					300							
Ala	Ser	His	Ser	Ala	Asn	Ser	Gln	Pro	Ser	Ala	Thr	Thr	Val	Thr	Ala			
305					310					315					320			
Ile	Thr	Pro	Ala	Pro	Ile	Gln	Gln	Pro	Val	Lys	Ser	Met	Arg	Val	Leu			
				325					330					335				
Lys	Pro	Glu	Leu	Gln	Thr	Ala	Leu	Ala	Pro	Thr	His	Pro	Ser	Trp	Met			
			340					345					350					
Pro	Gln	Pro	Val	Gln	Thr	Val	Gln	Pro	Thr	Pro	Phe	Ser	Glu	Gly	Thr			
		355					360					365						
Ala	Ser	Ser	Val	Pro	Val	Ile	Pro	Pro	Val	Ala	Glu	Ala	Pro	Ser	Tyr			
		370				375					380							
Gln	Gly	Pro	Pro	Pro	Pro	Tyr	Pro	Lys	His	Leu	Leu	His	Gln	Asn	Pro			
385					390					395					400			
Ser	Val	Pro	Pro	Tyr	Glu	Ser	Val	Ser	Lys	Pro	Cys	Lys	Asp	Glu	Gln			
				405					410					415				
Pro	Ser	Leu	Pro	Lys	Glu	Asp	Asp	S r	Glu	Lys	Ser	Ala	Asp	Ser	Gly			
			420					425					430					

Asp Ser ly Asp Lys Glu Lys Lys Gln Il Thr Thr Ser Pro Ile Thr  
 435 440 445  
 Val Arg Lys Asn Lys Lys Asp Glu Glu Arg Arg Glu Ser Arg Ile Gln  
 450 455 460  
 Ser Tyr Ser Pro Gln Ala Phe Lys Phe Phe Met Glu Gln His Val Glu  
 465 470 475 480  
~~Asn Val Leu Lys Ser His Gln Gln Arg Leu His Arg Lys Lys Gln Leu~~  
~~485 490 495~~  
 Glu Asn Glu Met Met Arg Val Gly Leu Ser Gln Asp Ala Gln Asp Gln  
 500 505 510  
 Met Arg Lys Met Leu Cys Gln Lys Glu Ser Asn Tyr Ile Arg Leu Lys  
 515 520 525  
 Arg Ala Lys Met Asp Lys Ser Met Phe Val Lys Ile Lys Thr Leu Gly  
 530 535 540  
 Ile Gly Ala Phe Gly Glu Val Cys Leu Ala Arg Lys Val Asp Thr Lys  
 545 550 555 560  
 Ala Leu Tyr Ala Thr Lys Thr Leu Arg Lys Lys Asp Val Leu Leu Arg  
 565 570 575  
 Asn Gln Val Ala His Val Lys Ala Glu Arg Asp Ile Leu Ala Glu Ala  
 580 585 590  
 Asp Asn Glu Trp Val Val Arg Leu Tyr Tyr Ser Phe Gln Asp Lys Asp  
 595 600 605  
 Asn Leu Tyr Phe Val Met Asp Tyr Ile Pro Gly Gly Asp Met Met Ser  
 610 615 620  
 Leu Leu Ile Arg Met Gly Ile Phe Pro Glu Asn Leu Ala Arg Phe Tyr  
 625 630 635 640  
 Ile Ala Glu Leu Thr Cys Ala Val Glu Ser Val His Lys Met Gly Phe  
 645 650 655  
 Ile His Arg Asp Ile Lys Pro Asp Asn Ile Leu Ile Asp Arg Asp Gly  
 660 665 670  
 His Ile Lys Leu Thr Asp Phe Gly Leu Cys Thr Gly Phe Arg Trp Thr  
 675 680 685  
 His Asp Ser Lys Tyr Tyr Gln Ser Gly Asp His Pro Arg Gln Asp Ser  
 690 695 700  
 Met Asp Phe Ser Asn Glu Trp Gly Asp Pro Ser Asn Cys Arg Cys Gly  
 705 710 715 720  
 Asp Arg Leu Lys Pro Leu Glu Arg Arg Ala Ala Arg Gln His Gln Arg  
 725 730 735  
 Cys Leu Ala His Ser Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu  
 740 745 750  
 Val Leu Leu Arg Thr Gly Tyr Thr Gln Leu Cys Asp Trp Trp Ser Val  
 755 760 765  
 Gly Val Ile Leu Cys Glu Met Leu Val Gly Gln Pro Pro Phe Leu Ala  
 770 775 780  
 Gln Thr Pro L u Glu Thr Gln Met Lys Val Ile Ile Trp Gln Thr Ser

785						790						795						800
L u His Ile Pro	Pro	Gln	Ala	Lys	Leu	Ser	Pro	Glu	Ala	Ser	Asp	Leu						
	805					810					815							
Il Ile Lys Leu	Cys	Arg	Gly	Pro	Glu	Asp	Arg	Leu	Gly	Lys	Asn	Gly						
	820				825					830								
Ala Asp Glu Ile	Lys	Ala	His	Pro	Phe	Phe	Lys	Thr	Ile	Asp	Phe	Ser						
	835			840					845									
Ser Asp Leu Arg	Gln	Gln	Ser	Ala	Ser	Tyr	Ile	Pro	Lys	Ile	Thr	His						
	850			855				860										
Pro Thr Asp Thr	Ser	Asn	Phe	Asp	Pro	Val	Asp	Pro	Asp	Lys	Leu	Trp						
	865		870				875					880						
Ser Asp Gly Ser	Glu	Glu	Glu	Asn	Ile	Ser	Asp	Thr	Leu	Ser	Gly	Trp						
	885					890					895							
Tyr Lys Asn Gly	Lys	His	Pro	Glu	His	Ala	Phe	Tyr	Glu	Phe	Thr	Phe						
	900				905				910									
Arg Arg Phe Phe	Asp	Asp	Asn	Gly	Tyr	Pro	Tyr	Asn	Tyr	Pro	Lys	Pro						
	915			920					925									
Ile Glu Tyr Glu	Tyr	Ile	His	Ser	Gln	Gly	Ser	Glu	Gln	Gln	Ser	Asp						
	930		935					940										
Glu Asp Asp Gln	His	Thr	Ser	Ser	Asp	Gly	Asn	Asn	Arg	Asp	Leu	Val						
	945		950				955				960							
Tyr Val *																		

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3155 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..2943

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATG AGA GCC ACC CCG AAG TTT GGA CCT TAT CAA AAA GCT CTC AGG GAA	48
Met Arg Ala Thr Pro Lys Phe Gly Pro Tyr Gln Lys Ala Leu Arg Glu	
1 5 10 15	
ATC CGA TAT TCC CTC CTG CCT TTT GCC AAC GAG TCA GGC ACT TCG GCA	96
Ile Arg Tyr Ser Leu Leu Pro Phe Ala Asn Glu Ser Gly Thr Ser Ala	
20 25 30	
GCT GCA GAG GTG AAC CGG CAG ATG CTT CAG GAG TTG GTG AAT GCG GCA	144
Ala Ala Glu Val Asn Arg Gln Met Leu Gln Glu Leu Val Asn Ala Ala	
35 40 45	
TGT GAC CAG GAG ATG GCT GGC AGA GCG CTC ACG CAG ACG GGC AGT AGG	192
Cys Asp Gln Glu Met Ala Gly Arg Ala Leu Thr Gln Thr Gly Ser Arg	

50					55					60						
AGT Ser 65	ATC Ile	GAA Glu	GCT Ala	GCC Ala	TTG Leu 70	GAG Glu	TAC Tyr	ATC Ile	AGT Ser	AAG Lys 75	ATG Met	GGC Gly	TAC Tyr	CTG L u	GAC Asp 80	240
CCC Pro	AGG Arg	AAT Asn	GAG Glu	CAG Gln 85	ATT Ile	GTG Val	CGA Arg	GTC Val	ATC Ile 90	AAG Lys	CAG Gln	ACC Thr	TCC Ser	CCA Pro 95	GGA Gly	288
AAG Lys	GGC Gly	CTG Leu	GCG Ala 100	TCC Ser	ACC Thr	CCG Pro	GTG Val	ACT Thr 105	CGG Arg	CGG Arg	CCC Pro	AGT Ser	TTC Phe 110	GAG Glu	GGC Gly	336
ACA Thr	GGG Gly	GAA Glu 115	GCA Ala	CTC Leu	CCA Pro	TCC Ser	TAC Tyr 120	CAC His	CAG Gln	CTG Leu	GGT Gly 125	GGT Gly	GCA Ala	AAC Asn	TAC Tyr	384
GAG Glu	GGC Gly 130	CCC Pro	GCC Ala	GCA Ala	CTG Leu	GAG Glu 135	GAG Glu	ATG Met	CCG Pro	CGG Arg	CAA Gln 140	TAT Tyr	TTA Leu	GAC Asp	TTT Phe	432
CTC Leu 145	TTC Phe	CCT Pro	GGA Gly	GCC Ala	GGA Gly 150	GCC Ala	GGC Gly	ACC Thr	CAC His	GGT Gly 155	GCC Ala	CAG Gln	GCT Ala	CAC His	CAG Gln 160	480
CAT His	CCT Pro	CCC Pro	AAA Lys	GGG Gly 165	TAC Tyr	AGC Ser	ACA Thr	GCA Ala	GTA Val 170	GAG Glu	CCA Pro	AGT Ser	GCG Ala	CAC His 175	TTT Phe	528
CCG Pro	GGC Gly	ACA Thr 180	CAC His	TAT Tyr	GGT Gly	CGT Arg	GGT Gly	CAT His 185	CTA Leu	CTA Leu	TCG Ser	GAG Glu	CAG Gln 190	TCT Ser	GGG Gly	576
TAT Tyr	GGG Gly	GTG Val 195	CAG Gln	CGC Arg	AGT Ser	TCC Ser	TCC Ser 200	TTC Phe	CAG Gln	AAC Asn	AAG Lys	ACG Thr 205	CCA Pro	CCA Pro	GAT Asp	624
GCC Ala	TAT Tyr 210	TCC Ser	AGC Ser	ATG Met	GCC Ala	AAG Lys 215	GCC Ala	CAG Gln	GGT Gly	GGC Gly	CCT Pro 220	CCC Pro	GCC Ala	AGC Ser	CTC Leu	672
ACC Thr 225	TTT Phe	CCT Pro	GCC Ala	CAT His	GCT Ala 230	GGG Gly	CTG Leu	TAC Tyr	ACT Thr	GCC Ala 235	TCG Ser	CAC His	CAC His	AAG Lys	CCG Pro 240	720
GCG Ala	GCT Ala	ACC Thr	CCA Pro	CCT Pro 245	GGG Gly	GCC Ala	CAC His	CCA Pro	TTA Leu 250	CAT His	GTG Val	TTG Leu	GGC Gly	ACC Thr 255	CGG Arg	768
GGT Gly	CCC Pro	ACG Thr	TTT Phe 260	ACT Thr	GGC Gly	GAA Glu	AGC Ser	TCT Ser 265	GCA Ala	CAG Gln	GCT Ala	GTG Val	CTG Leu 270	GCA Ala	CCG Pro	816
TCC Ser	AGG Arg	AAC Asn 275	AGC Ser	CTC Leu	AAT Asn	GCT Ala	GAC Asp 280	TTG Leu	TAC Tyr	GAG Glu	CTG Leu	GGC Gly 285	TCC Ser	ACG Thr	GTG Val	864
CCC Pro	TGG Trp 290	TCT Ser	GCA Ala	GCT Ala	CCA Pro	CTG Leu 295	GCA Ala	CGC Arg	CGC Arg	GAC Asp	TCG Ser 300	CTG Leu	CAG Gln	AAG Lys	CAG Gln	912
GGT Gly 305	CTA Leu	GAA Glu	GCC Ala	TCG Ser	CGG Arg 310	CCG Pro	CAT His	GTG Val	GCT Ala	TTT Phe 315	CGG Arg	GCT Ala	GGC Gly	CCC Pro	AGC Ser 320	960
AGG	ACC	AAC	TCC	TTC	AAC	AAC	CCA	CAA	CCT	GAG	CCC	TCA	CTG	CCC	GCC	1008

Arg	Thr	Asn	Ser	Ph	Asn	Asn	Pro	Gln	Pr	Glu	Pro	Ser	Leu	Pro	Ala	
				325					330					335		
CCC	AAC	ACG	GTC	ACC	GCC	GTG	ACG	GCC	CA	CAC	ATC	CTT	CAC	CCT	GTG	1056
Pro	Asn	Thr	Val	Thr	Ala	Val	Thr	Ala	Ala	His	Ile	L u	His	Pro	Val	
			340					345					350			
AAG	AGC	GTG	CGT	GTG	CTG	CGG	CCC	GAG	CCC	CAG	ACA	GCC	GTG	GGG	CCC	1104
Lys	Ser	Val	Arg	Val	Leu	Arg	Pro	Glu	Pro	Gln	Thr	Ala	Val	Gly	Pro	
		355					360					365				
TCG	CAC	CCC	GCC	TGG	GTG	GCT	GCG	CCC	ACA	GCA	CCT	GCC	ACT	GAG	AGC	1152
Ser	His	Pro	Ala	Trp	Val	Ala	Ala	Pro	Thr	Ala	Pro	Ala	Thr	Glu	Ser	
	370					375					380					
CTG	GAG	ACG	AAG	GAG	GGC	AGC	GCA	GGC	CCA	CAC	CCG	CTG	GAT	GTG	GAC	1200
Leu	Glu	Thr	Lys	Glu	Gly	Ser	Ala	Gly	Pro	His	Pro	Leu	Asp	Val	Asp	
	385				390					395					400	
TAT	GGC	GGC	TCC	GAG	CGC	AGG	TGC	CCA	CCG	CCT	CCG	TAT	CCA	AAG	CAC	1248
Tyr	Gly	Gly	Ser	Glu	Arg	Arg	Cys	Pro	Pro	Pro	Pro	Tyr	Pro	Lys	His	
			405						410					415		
TTG	CTG	CTG	CCC	AGT	AAG	TCT	GAG	CAG	TAC	AGC	GTG	GAC	CTG	GAC	AGC	1296
Leu	Leu	Leu	Pro	Ser	Lys	Ser	Glu	Gln	Tyr	Ser	Val	Asp	Leu	Asp	Ser	
			420					425					430			
CTG	TGC	ACC	AGT	GTG	CAG	CAG	AGT	CTG	CGA	GGG	GGC	ACT	GAT	CTA	GAC	1344
Leu	Cys	Thr	Ser	Val	Gln	Gln	Ser	Leu	Arg	Gly	Gly	Thr	Asp	Leu	Asp	
		435					440					445				
GGG	AGT	GAC	AAG	AGC	CAC	AAA	GGT	GCG	AAG	GGA	GAC	AAA	GCT	GGC	AGA	1392
Gly	Ser	Asp	Lys	Ser	His	Lys	Gly	Ala	Lys	Gly	Asp	Lys	Ala	Gly	Arg	
	450					455					460					
GAC	AAA	AAG	CAG	ATT	CAG	ACC	TCC	CCG	GTG	CCT	GTC	CGC	AAG	AAT	AGC	1440
Asp	Lys	Lys	Gln	Ile	Gln	Thr	Ser	Pro	Val	Pro	Val	Arg	Lys	Asn	Ser	
	465				470					475					480	
AGA	GAT	GAA	GAG	AAG	AGA	GAG	TCT	CGC	ATC	AAG	AGT	TAC	TCC	CCT	TAT	1488
Arg	Asp	Glu	Glu	Lys	Arg	Glu	Ser	Arg	Ile	Lys	Ser	Tyr	Ser	Pro	Tyr	
				485					490					495		
GCC	TTC	AAA	TTC	TTC	ATG	GAG	CAA	CAC	GTG	GAG	AAT	GTC	ATC	AAA	ACC	1536
Ala	Phe	Lys	Phe	Phe	Met	Glu	Gln	His	Val	Glu	Asn	Val	Ile	Lys	Thr	
			500					505					510			
TAC	CAG	CAG	AAG	GTC	AGC	CGG	AGG	CTA	CAG	CTG	GAG	CAG	GAA	ATG	GCC	1584
Tyr	Gln	Gln	Lys	Val	Ser	Arg	Arg	Leu	Gln	Leu	Glu	Gln	Glu	Met	Ala	
		515					520					525				
AAA	GCT	GGG	CTC	TGT	GAG	GCC	GAG	CAG	GAG	CAG	ATG	AGG	AAG	ATC	CTC	1632
Lys	Ala	Gly	Leu	Cys	Glu	Ala	Glu	Gln	Glu	Gln	Met	Arg	Lys	Ile	Leu	
	530					535					540					
TAC	CAG	AAG	GAG	TCT	AAC	TAC	AAC	CGG	CTG	AAG	AGG	GCC	AAG	ATG	GAC	1680
Tyr	Gln	Lys	Glu	Ser	Asn	Tyr	Asn	Arg	Leu	Lys	Arg	Ala	Lys	Met	Asp	
	545				550					555				560		
AAG	TCC	ATG	TTT	GTG	AAA	ATC	AAG	ACT	CTA	GGC	ATC	GGT	GCC	TTT	GGG	1728
Lys	Ser	Met	Phe	Val	Lys	Ile	Lys	Thr	Leu	Gly	Ile	Gly	Ala	Phe	Gly	
				565					570					575		
GAA	GTG	TGC	CTC	GCT	TGT	AAG	CTG	GAC	ACT	CAC	GCT	CTG	TAC	GCC	ATG	1776
Glu	Val	Cys	Leu	Ala	Cys	Lys	Leu	Asp	Thr	His	Ala	Leu	Tyr	Ala	M t	
			580					585					590			



AAG Lys	ACT Thr	CTC Leu	AGG Arg	AAG Lys	AAG Lys	GAT Asp	GTC Val	CTG L u	AAC Asn	CGG Arg	AAT Asn	CAA Gln	GTG Val	GCC Ala	CAT His	1824
		595					600					605				
GTC Val	AAG Lys	GCT Ala	GAG Glu	AGG Arg	GAC Asp	ATC Ile	CTG Leu	GCT Ala	GAA Glu	GCA Ala	GAC Asp	AAT Asn	GAG Glu	TGG Trp	GTG Val	1872
	610					615					620					
GTC Val	AAA Lys	CTC Leu	TAC Tyr	TAC Tyr	TCC Ser	TTC Phe	CAG Gln	GAC Asp	AAG Lys	GAC Asp	AGC Ser	CTG Leu	TAC Tyr	TTT Phe	GTG Val	1920
	625				630					635					640	
ATG Met	GAC Asp	TAC Tyr	ATA Ile	CCA Pro	GGC Gly	GGG Gly	GAT Asp	ATG Met	ATG Met	AGC Ser	CTG Leu	CTG Leu	ATC Ile	AGG Arg	ATG Met	1968
				645					650					655		
GAG Glu	GTC Val	TTC Phe	CCT Pro	GAG Glu	CAC His	CTG Leu	GCC Ala	CGC Arg	TTC Phe	TAC Tyr	ATT Ile	GCA Ala	GAG Glu	TTG Leu	ACC Thr	2016
			660					665					670			
CTG Leu	GCC Ala	ATT Ile	GAA Glu	AGT Ser	GTC Val	CAC His	AAG Lys	ATG Met	GGC Gly	TTT Phe	ATC Ile	CAC His	CGG Arg	GAC Asp	ATC Ile	2064
		675					680					685				
AAG Lys	CCT Pro	GAC Asp	AAC Asn	ATA Ile	CTC Leu	ATC Ile	GAC Asp	CTG Leu	GAT Asp	GGT Gly	CAT His	ATT Ile	AAG Lys	CTG Leu	ACA Thr	2112
	690					695					700					
GAT Asp	TTT Phe	GGC Gly	CTC Leu	TGC Cys	ACT Thr	GGA Gly	TTC Phe	AGG Arg	TGG Trp	ACT Thr	CAC His	AAT Asn	TCC Ser	AAG Lys	TAC Tyr	2160
	705				710					715					720	
TAC Tyr	CAG Gln	AAA Lys	GGG Gly	AAC Asn	CAC His	ATG Met	AGA Arg	CAG Gln	GAC Asp	AGC Ser	ATG Met	GAG Glu	CCC Pro	GGT Gly	GAC Asp	2208
				725					730					735		
CTC Leu	TGG Trp	GAC Asp	GAT Asp	GTT Val	TCC Ser	AAC Asn	TGT Cys	CGC Arg	TGT Cys	GGA Gly	GAC Asp	AGG Arg	TTA Leu	AAG Lys	ACC Thr	2256
			740					745					750			
CTG Leu	GAG Glu	CAG Gln	AGG Arg	GCG Ala	CAG Gln	AAG Lys	CAG Gln	CAC His	CAG Gln	AGG Arg	TGC Cys	CTG Leu	GCA Ala	CAT His	TCT Ser	2304
		755					760					765				
CTT Leu	GTC Val	GGG Gly	ACA Thr	CCA Pro	AAT Asn	TAC Tyr	ATC Ile	GCT Ala	CCG Pro	GAG Glu	GTG Val	CTT Leu	CTC Leu	CGC Arg	AAA Lys	2352
	770					775					780					
GGG Gly	TAC Tyr	ACG Thr	CAG Gln	CTC Leu	TGT Cys	GAC Asp	TGG Trp	TGG Trp	AGC Ser	GTC Val	GGT Gly	GTG Val	ATT Ile	CTC Leu	TTT Phe	2400
	785				790					795					800	
GAG Glu	ATG Met	CTG Leu	GTT Val	GGG Gly	CAG Gln	CCG Pro	CCT Pro	TTC Phe	TTG Leu	GCC Ala	CCC Pro	ACC Thr	CCC Pro	ACA Thr	GAG Glu	2448
				805					810					815		
ACG Thr	CAG Gln	CTG Leu	AAG Lys	GTG Val	ATC Ile	AAC Asn	TGG Trp	GAG Glu	AGC Ser	ACG Thr	CTG Leu	CAT His	ATC Ile	CCT Pro	ACG Thr	2496
			820					825					830			
CAG Gln	GTG Val	AGG Arg	CTC Leu	AGC Ser	GCT Ala	GAG Glu	GCC Ala	CGA Arg	GAC Asp	CTC Leu	ATC Ile	ACG Thr	AAG Lys	CTG Leu	TGC Cys	2544
		835					840					845				
TGC Cys	GCG Ala	GCT Ala	GAC Asp	TGC Cys	CGC Arg	CTG Leu	GGC Gly	AGG Arg	GAT Asp	GGG Gly	GCA Ala	GAT Asp	GAC Asp	CTC L u	AAG Lys	2592
	850					855					860					

GCA CAC CCG TTC TTC AAC ACC ATC GAC TTT TCC CGT GAC ATC CGA AAG Ala His Pro Phe Phe Asn Thr Ile Asp Phe S r Arg Asp Ile Arg Lys 865 870 875 880	2640
CAG GCT GCA CCC TAC GTC CCC ACC ATC A C CAC CCC ATG GAC ACC TCC Gln Ala Ala Pro Tyr Val Pro Thr Ile Ser His Pro Met Asp Thr Ser 885 890 895	2688
AAT TTT GAC CCG GTG GAT GAA GAA AGC CCC TGG CAC GAG GCC AGC GGA Asn Phe Asp Pro Val Asp Glu Glu Ser Pro Trp His Glu Ala Ser Gly 900 905 910	2736
GAG AGC GCC AAG GCC TGG GAC ACG CTG GCC TCC CCC AGC AGC AAG CAT Glu Ser Ala Lys Ala Trp Asp Thr Leu Ala Ser Pro Ser Ser Lys His 915 920 925	2784
CCA GAG CAC GCC TTC TAT GAG TTC ACC TTC CGC AGG TTC TTC GAT GAC Pro Glu His Ala Phe Tyr Glu Phe Thr Phe Arg Arg Phe Phe Asp Asp 930 935 940	2832
AAC GGC TAT CCC TTC CGG TGC CCG AAG CCC TCA GAG CCC GCA GAG AGT Asn Gly Tyr Pro Phe Arg Cys Pro Lys Pro Ser Glu Pro Ala Glu Ser 945 950 955 960	2880
GCA GAC CCA GGG GAT GCG GAC TTG GAA GGT GCG GCC GAG GGC TGC CAG Ala Asp Pro Gly Asp Ala Asp Leu Glu Gly Ala Ala Glu Gly Cys Gln 965 970 975	2928
CCG GTG TAC GTG TAA GCCTCAGTTA ACCACAACCTC GAGGAAACCC AAAATGAGAT Pro Val Tyr Val *	2983
TTCTTTTCAG AAGACAAACT CAAGCTTAGG AATCCTTCAT TTTAGTTCT GGTAAATGGG	3043
CAACAGGAAG AGTCAACATG ATTTCAAATT AGCCCTCTGA GGACCTTCAC TGCATTAAAA	3103
CAGTATTTTT TAAAAAATTA GTACAGTATG GAAAGAGCAC TTATTTTGGG GG	3155

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 981 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Arg Ala Thr Pro Lys Phe Gly Pro Tyr Gln Lys Ala Leu Arg Glu  
1 5 10 15

Ile Arg Tyr Ser Leu Leu Pro Phe Ala Asn Glu Ser Gly Thr Ser Ala  
20 25 30

Ala Ala Glu Val Asn Arg Gln Met Leu Gln Glu Leu Val Asn Ala Ala  
35 40 45

Cys Asp Gln Glu Met Ala Gly Arg Ala Leu Thr Gln Thr Gly Ser Arg  
50 55 60

Ser Ile Glu Ala Ala Leu Glu Tyr Ile Ser Lys M t Gly Tyr Leu Asp  
65 70 75 80

Pro Arg Asn Glu Gln Il Val Arg Val Ile Lys Gln Thr Ser Pro Gly

85										90				95			
Lys	Gly	Leu	Ala	Ser	Thr	Pro	Val	Thr	Arg	Arg	Pro	Ser	Phe	Glu	Gly		
			100					105					110				
Thr	Gly	Glu	Ala	Leu	Pro	Ser	Tyr	His	Gln	Leu	Gly	Gly	Ala	Asn	Tyr		
		115					120					125					
Glu	Gly	Pro	Ala	Ala	Leu	Glu	Glu	Met	Pro	Arg	Gln	Tyr	Leu	Asp	Phe		
	130					135					140						
Leu	Phe	Pro	Gly	Ala	Gly	Ala	Gly	Thr	His	Gly	Ala	Gln	Ala	His	Gln		
145					150					155					160		
His	Pro	Pro	Lys	Gly	Tyr	Ser	Thr	Ala	Val	Glu	Pro	Ser	Ala	His	Phe		
				165					170					175			
Pro	Gly	Thr	His	Tyr	Gly	Arg	Gly	His	Leu	Leu	Ser	Glu	Gln	Ser	Gly		
			180					185					190				
Tyr	Gly	Val	Gln	Arg	Ser	Ser	Ser	Phe	Gln	Asn	Lys	Thr	Pro	Pro	Asp		
		195					200					205					
Ala	Tyr	Ser	Ser	Met	Ala	Lys	Ala	Gln	Gly	Gly	Pro	Pro	Ala	Ser	Leu		
	210					215					220						
Thr	Phe	Pro	Ala	His	Ala	Gly	Leu	Tyr	Thr	Ala	Ser	His	His	Lys	Pro		
225					230					235					240		
Ala	Ala	Thr	Pro	Pro	Gly	Ala	His	Pro	Leu	His	Val	Leu	Gly	Thr	Arg		
				245					250					255			
Gly	Pro	Thr	Phe	Thr	Gly	Glu	Ser	Ser	Ala	Gln	Ala	Val	Leu	Ala	Pro		
			260					265					270				
Ser	Arg	Asn	Ser	Leu	Asn	Ala	Asp	Leu	Tyr	Glu	Leu	Gly	Ser	Thr	Val		
		275					280					285					
Pro	Trp	Ser	Ala	Ala	Pro	Leu	Ala	Arg	Arg	Asp	Ser	Leu	Gln	Lys	Gln		
	290					295					300						
Gly	Leu	Glu	Ala	Ser	Arg	Pro	His	Val	Ala	Phe	Arg	Ala	Gly	Pro	Ser		
305					310					315					320		
Arg	Thr	Asn	Ser	Phe	Asn	Asn	Pro	Gln	Pro	Glu	Pro	Ser	Leu	Pro	Ala		
				325					330					335			
Pro	Asn	Thr	Val	Thr	Ala	Val	Thr	Ala	Ala	His	Ile	Leu	His	Pro	Val		
			340					345					350				
Lys	Ser	Val	Arg	Val	Leu	Arg	Pro	Glu	Pro	Gln	Thr	Ala	Val	Gly	Pro		
		355					360					365					
Ser	His	Pro	Ala	Trp	Val	Ala	Ala	Pro	Thr	Ala	Pro	Ala	Thr	Glu	Ser		
	370					375					380						
Leu	Glu	Thr	Lys	Glu	Gly	Ser	Ala	Gly	Pro	His	Pro	Leu	Asp	Val	Asp		
385					390					395					400		
Tyr	Gly	Gly	Ser	Glu	Arg	Arg	Cys	Pro	Pro	Pro	Pro	Tyr	Pro	Lys	His		
				405					410					415			
Leu	Leu	Leu	Pro	Ser	Lys	Ser	Glu	Gln	Tyr	Ser	Val	Asp	Leu	Asp	Ser		
			420														

Gly Ser Asp Lys Ser His Lys Gly Ala Lys Gly Asp Lys Ala Gly Arg  
 450 455 460  
 Asp Lys Lys Gln Il Gln Thr S r Pro Val Pr Val Arg Lys Asn Ser  
 465 470 475 480  
 Arg Asp Glu Glu Lys Arg Glu Ser Arg Ile Lys Ser Tyr Ser Pro Tyr  
 485 490 495  
~~Ala Phe Lys Phe Phe Met Glu Gln His Val Glu Asn Val Ile Lys Thr~~  
~~500 505 510~~  
 Tyr Gln Gln Lys Val Ser Arg Arg Leu Gln Leu Glu Gln Glu Met Ala  
 515 520 525  
 Lys Ala Gly Leu Cys Glu Ala Glu Gln Glu Gln Met Arg Lys Ile Leu  
 530 535 540  
 Tyr Gln Lys Glu Ser Asn Tyr Asn Arg Leu Lys Arg Ala Lys Met Asp  
 545 550 555 560  
 Lys Ser Met Phe Val Lys Ile Lys Thr Leu Gly Ile Gly Ala Phe Gly  
 565 570 575  
 Glu Val Cys Leu Ala Cys Lys Leu Asp Thr His Ala Leu Tyr Ala Met  
 580 585 590  
 Lys Thr Leu Arg Lys Lys Asp Val Leu Asn Arg Asn Gln Val Ala His  
 595 600 605  
 Val Lys Ala Glu Arg Asp Ile Leu Ala Glu Ala Asp Asn Glu Trp Val  
 610 615 620  
 Val Lys Leu Tyr Tyr Ser Phe Gln Asp Lys Asp Ser Leu Tyr Phe Val  
 625 630 635 640  
 Met Asp Tyr Ile Pro Gly Gly Asp Met Met Ser Leu Leu Ile Arg Met  
 645 650 655  
 Glu Val Phe Pro Glu His Leu Ala Arg Phe Tyr Ile Ala Glu Leu Thr  
 660 665 670  
 Leu Ala Ile Glu Ser Val His Lys Met Gly Phe Ile His Arg Asp Ile  
 675 680 685  
 Lys Pro Asp Asn Ile Leu Ile Asp Leu Asp Gly His Ile Lys Leu Thr  
 690 695 700  
 Asp Phe Gly Leu Cys Thr Gly Phe Arg Trp Thr His Asn Ser Lys Tyr  
 705 710 715 720  
 Tyr Gln Lys Gly Asn His Met Arg Gln Asp Ser Met Glu Pro Gly Asp  
 725 730 735  
 Leu Trp Asp Asp Val Ser Asn Cys Arg Cys Gly Asp Arg Leu Lys Thr  
 740 745 750  
 Leu Glu Gln Arg Ala Gln Lys Gln His Gln Arg Cys Leu Ala His Ser  
 755 760 765  
 Leu Val Gly Thr Pro Asn Tyr Ile Ala Pro Glu Val Leu Leu Arg Lys  
 770 775 780  
 Gly Tyr Thr Gln Leu Cys Asp Trp Trp Ser Val Gly Val Ile L u Phe  
 785 790 795 800  
 Glu Met Leu Val Gly Gln Pro Pro Phe Leu Ala Pro Thr Pro Thr Glu

[illegible]

(2) INFORMATION FOR SEQ ID NO:9:

**(i) SEQUENCE CHARACTERISTICS:**

- SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Asp Leu Lys Pro Glu Asn  
1 5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide  
(B) LOCATION: 2  
(D) OTHER INFORMATION: /label= A

/note= "X at the 8 cond position can be either Threonin or

**Serine."**

**(ix) FEATURE:**

(A) NAME/KEY: Peptid

(B) LOCATION: 5

(D) OTHER INFORMATION: /label= B

/note= "X at the fifth position can either be Tyrosine or Phenylalanine."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gly Xaa Xaa Xaa Xaa Xaa Ala Pro Glu  
1 5

**(2) INFORMATION FOR SEQ ID NO:11:**

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 620 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Asp Asn Thr Asn Arg Pro His Leu Asn Leu Gly Thr Asn Asp Thr  
1 5 10 15

Arg Met Ala Pro Asn Asp Arg Thr Tyr Pro Thr Thr Pro Ser Thr Phe  
20 25 30

Pro Gln Pro Val Phe Pro Gly Gln Gln Ala Gly Gly Ser Gln Gln Tyr  
35 40 45

Asn Gln Ala Tyr Ala Gln Ser Gly Asn Tyr Tyr Gln Gln Asn His Asn  
50 55 60

Asp Pro Asn Thr Gly Leu Ala His Gln Phe Ala His Gln Asn Ile Gly  
65 70 75 80

Ser Ala Gly Arg Ala Ser Pro Tyr Gly Ser Arg Gly Pro Ser Pro Ala  
85 90 95

Gln Arg Pro Arg Thr Ser Gly Asn Ser Gly Gln Gln Gln Thr Tyr Gly  
100 105 110

Asn Tyr Leu Ser Ala Pro Met Pro Ser Asn Thr Gln Thr Glu Phe Ala  
115 120 125

Pro Leu Pro Ser Gly Thr Pro Thr Asn Met Ala Pro Met Pro Thr Thr  
130 135 140

Thr Arg Arg Ser Ala His Ser Trp Pro Leu Thr Ser Leu Arg Thr Ala  
145 150 155 160

Ser Ser Ala Pro Gly Ser Ala Thr Arg Gly Glu Cys Cys Ser Asp Ala  
165 170 175

Leu Leu Pro Leu His Pro Ala Val Ile Gly Ala Asp Thr Leu Phe Arg  
180 185 190

Gln Ser Glu Met Glu Gln Lys Leu Gly Glu Thr Asn Asp Ala Arg Arg  
195 200 205

Arg Glu Ser Ile Trp Ser Thr Ala Gly Arg Lys Glu Gly Gln Tyr Leu  
 210 215 220  
 Arg Phe Leu Arg Thr Lys Asp Lys Pro Glu Asn Tyr Gln Thr Ile Lys  
 225 230 235 240  
 Ile Ile Gly Lys Gly Ala Phe Gly Glu Val Lys L u Val Gln Lys Lys  
 245 250 255  
~~Ala Asp Gly Lys Val Tyr Ala Met Lys Ser Leu Ile Lys Thr Glu Met~~  
~~260 265 270~~  
 Phe Lys Lys Asp Gln Leu Ala His Val Arg Ala Glu Arg Asp Ile Leu  
 275 280 285  
 Ala Glu Ser Asp Ser Pro Trp Val Val Lys Leu Tyr Thr Thr Phe Gln  
 290 295 300  
 Asp Ala Asn Phe Leu Tyr Met Leu Met Glu Phe Leu Pro Gly Gly Asp  
 305 310 315 320  
 Leu Met Thr Met Leu Ile Lys Tyr Glu Ile Phe Ser Glu Asp Ile Thr  
 325 330 335  
 Arg Phe Tyr Ile Ala Glu Ile Val Leu Ala Ile Asp Ala Val His Lys  
 340 345 350  
 Leu Gly Phe Ile His Arg Asp Ile Lys Pro Asp Asn Ile Leu Leu Asp  
 355 360 365  
 Arg Gly Gly His Val Lys Leu Thr Asp Phe Gly Leu Ser Thr Gly Phe  
 370 375 380  
 His Lys Leu His Asp Asn Asn Tyr Tyr Thr Gln Leu Leu Gln Gly Lys  
 385 390 395 400  
 Ser Asn Lys Pro Arg Asp Asn Arg Asn Ser Val Ala Ile Asp Gln Ile  
 405 410 415  
 Asn Leu Thr Val Ser Asn Arg Ala Gln Ile Asn Asp Trp Arg Arg Ser  
 420 425 430  
 Arg Arg Leu Met Ala Tyr Ser Thr Val Gly Thr Pro Asp Tyr Ile Ala  
 435 440 445  
 Pro Glu Ile Phe Thr Gly His Gly Tyr Ser Phe Asp Cys Asp Trp Trp  
 450 455 460  
 Ser Leu Gly Thr Ile Met Phe Glu Cys Leu Val Gly Trp Pro Pro Phe  
 465 470 475 480  
 Cys Ala Glu Asp Ser His Asp Thr Tyr Arg Lys Ile Val Asn Trp Arg  
 485 490 495  
 His Ser Leu Tyr Phe Pro Asp Asp Ile Thr Leu Gly Val Asp Ala Glu  
 500 505 510  
 Asn Leu Ile Arg Ser Leu Ile Cys Asn Thr Glu Asn Arg Leu Gly Arg  
 515 520 525  
 Gly Gly Ala His Glu Ile Lys Ser His Ala Phe Phe Arg Gly Val Glu  
 530 535 540  
 Phe Asp Ser Leu Arg Arg Ile Arg Ala Pro Phe Glu Pro Arg Leu Thr  
 545 550 555 560  
 Ser Ala Ile Asp Thr Thr Tyr Phe Pro Thr Asp Glu Il Asp Gln Thr

	565		570		575
Asp Asn Ala Thr Leu Leu Lys Ala	580	In Gln Ala Ala Arg Gly Ala Ala	585	590	
Ala Pro Ala Gln Gln Glu Glu Ser	595	Pro Glu Leu Ser Leu Pro Phe Ile	600	605	
Gly Tyr Thr Phe Lys Arg Phe Asp Asn Asn Phe Arg	610		615	620	

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 526 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met	Asp	Ser	Ala	Arg	Gly	Trp	Phe	Gln	Lys	Leu	Ser	Ser	Thr	Lys	Lys	1	5	10	15
Asp	Pro	Met	Ala	Ser	Gly	Arg	Glu	Asp	Gly	Lys	Pro	Val	Ser	Ala	Glu	20	25	30	
Glu	Ala	Ser	Asn	Ile	Thr	Lys	Gln	Arg	Val	Ala	Ala	Ala	Lys	Gln	Tyr	35	40	45	
Ile	Glu	Lys	His	Tyr	Arg	Glu	Gln	Met	Lys	Asn	Leu	Gln	Glu	Arg	Arg	50	55	60	
Glu	Arg	Arg	Ile	Leu	Leu	Glu	Lys	Lys	Leu	Ala	Asp	Ala	Asp	Val	Ser	65	70	75	80
Glu	Glu	Asp	Gln	Asn	Asn	Leu	Leu	Lys	Phe	Leu	Glu	Lys	Lys	Glu	Thr	85	90	95	
Glu	Tyr	Met	Arg	Leu	Gln	Arg	His	Lys	Met	Gly	Ala	Asp	Asp	Phe	Glu	100	105	110	
Leu	Leu	Thr	Met	Ile	Gly	Lys	Gly	Ala	Phe	Gly	Glu	Pro	Ile	Cys	Met	115	120	125	
Ile	Gly	Phe	Ser	Val	Ile	Thr	Gly	Gln	Asn	Cys	Arg	Glu	Lys	Thr	Thr	130	135	140	
Gly	Gln	Val	Tyr	Ala	Met	Lys	Lys	Leu	Lys	Lys	Ser	Glu	Met	Leu	Arg	145	150	155	160
Arg	Gly	Gln	Val	Glu	His	Val	Lys	Ala	Glu	Arg	Asn	Leu	Leu	Ala	Glu	165	170	175	
Val	Asp	Ser	Asp	Cys	Ile	Val	Lys	Leu	Tyr	Tyr	Ser	Phe	Gln	Asp	Asp	180	185	190	
Asp	Tyr	Leu	Tyr	Leu	Val	Met	Glu	Tyr	Leu	Pro	Gly	Gly	Asp	Met	Met	195	200	205	
Thr	Leu	Leu	Met	Arg	Lys	Asp	Ile	Leu	Thr	Glu	Asp	Glu	Ala	Arg	Phe	210	215	220	



Tyr Val Ala Glu Thr Val Leu Ala Ile Glu Ser Ile His Lys His Asn  
 225 230 235 240  
 Tyr Ile His Arg Asp Ile Lys Pro Asp Asn Leu Leu L u Asp Arg Tyr  
 245 250 255  
 Gly His L u Lys Leu Ser Asp Phe Gly Leu Cys Lys Pro Leu Asp Cys  
 260 265 270  
 Ser Thr Leu Glu Glu Lys Asp Phe Ser Val Gly Asp Asn Ala Asn Gly  
 275 280 285  
 Gly Ser Arg Ser Asp Ser Pro Pro Ala Pro Lys Arg Thr Gln Gln Glu  
 290 295 300  
 Gln Leu Glu His Trp Gln Lys Asn Arg Arg Met Leu Ala Tyr Ser Thr  
 305 310 315 320  
 Val Gly Thr Pro Asp Tyr Ile Ala Pro Glu Val Leu Leu Lys Lys Gly  
 325 330 335  
 Tyr Gly Met Glu Cys Asp Trp Trp Ser Leu Gly Ala Ile Met Tyr Glu  
 340 345 350  
 Met Leu Val Gly Tyr Pro Pro Phe Tyr Ser Asp Asp Pro Met Ser Thr  
 355 360 365  
 Cys Arg Lys Ile Val Asn Trp Lys Asn His Leu Lys Phe Pro Glu Glu  
 370 375 380  
 Ala Lys Leu Ser Pro Glu Ala Lys Asp Ile Ile Ser Arg Leu Leu Cys  
 385 390 395 400  
 Asn Val Thr Glu Arg Leu Gly Ser Asn Gly Ala Asp Glu Ile Lys Val  
 405 410 415  
 His Ser Trp Phe Lys Gly Ile Asp Trp Asp Arg Ile Tyr Gln Met Glu  
 420 425 430  
 Ala Ala Phe Ile Pro Glu Val Asn Asp Glu Leu Asp Thr Gln Asn Phe  
 435 440 445  
 Glu Lys Phe Glu Glu Ser Glu Ser His Ser Gln Ser Gly Ser Arg Ser  
 450 455 460  
 Gly Pro Trp Arg Lys Met Leu Ser Ser Lys Asp Ile Asn Phe Val Gly  
 465 470 475 480  
 Tyr Thr Tyr Lys Asn Phe Lys Val Val Asn Asp Tyr Gln Val Pro Gly  
 485 490 495  
 Met Val Glu Leu Lys Lys Thr Asn Thr Lys Pro Lys Lys Pro Thr Ile  
 500 505 510  
 Lys Ser Leu Phe Gly Asp Glu Ser Glu Ala Ser Glu Asp Asn  
 515 520 525

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 479 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Arg Lys Leu His Asp Ala Asp Val S r Glu Glu Asp Gln Asn Asn Leu  
 1 5 10 15  
 Leu Lys Phe Leu Glu Lys Lys Glu Thr Glu Tyr Met Arg L u Gln Arg  
 20 25 30  
~~His Lys Met Gly Ala Asp Asp Phe Glu Leu Leu Thr Met Ile Gly Lys~~  
~~35 40 45~~  
 Gly Ala Phe Gly Glu Val Arg Val Cys Arg Glu Lys Thr Thr Gly His  
 50 55 60  
 Val Tyr Ala Met Lys Lys Leu Lys Lys Ser Glu Met Leu Arg Arg Gly  
 65 70 75 80  
 Gln Val Glu His Val Lys Ala Glu Arg Asn Leu Leu Ala Glu Val Asp  
 85 90 95  
 Ser Asn Cys Ile Val Lys Leu Tyr Cys Ser Phe Gln Asp Glu Glu Tyr  
 100 105 110  
 Leu Tyr Leu Ile Met Glu Tyr Leu Pro Gly Gly Asp Met Met Thr Leu  
 115 120 125  
 Leu Met Arg Lys Asp Thr Leu Thr Glu Asp Glu Ala Arg Phe Tyr Val  
 130 135 140  
 Ala Glu Thr Ile Leu Ala Ile Glu Ser Ile His Lys His Asn Tyr Ile  
 145 150 155 160  
 His Arg Asp Ile Lys Pro Asp Asn Leu Leu Leu Asp Lys Phe Gly His  
 165 170 175  
 Leu Arg Leu Ser Asp Phe Gly Leu Cys Lys Pro Leu Asp Cys Ser Thr  
 180 185 190  
 Leu Glu Glu Lys Asp Phe Glu Val Asn Asn Gly Asn Gly Gly Ser Pro  
 195 200 205  
 Ser Asn Glu Gly Ser Thr Lys Pro Arg Arg Thr Gln Gln Glu Gln Leu  
 210 215 220  
 Gln His Trp Gln Lys Asn Arg Arg Met Leu Ala Tyr Ser Thr Val Gly  
 225 230 235 240  
 Thr Pro Asp Tyr Ile Ala Pro Glu Val Leu Leu Lys Lys Gly Tyr Gly  
 245 250 255  
 Met Glu Cys Asp Trp Trp Ser Leu Gly Ala Ile Met Tyr Glu Met Leu  
 260 265 270  
 Val Gly Tyr Pro Pro Phe Tyr Ser Asp Asp Pro Met Ser Thr Cys Arg  
 275 280 285  
 Lys Ile Val Asn Trp Arg Thr His Leu Lys Phe Pro Glu Glu Ala Lys  
 290 295 300  
 Leu Ser Pro Glu Ala Lys Asp Leu Ile Ser Lys Leu Leu Cys Asn Val  
 305 310 315 320  
 Thr Gln Arg Leu Gly Ser Asn Gly Ala His Glu Ile Lys Leu His Pro  
 325 330 335  
 Trp Phe Asn Gly Ile Asp Trp Glu Arg Ile Tyr Gln Met Glu Ala Ala

	340		345		350
Ph	Ile Pro Glu Val Asn Asp	Glu Leu Asp Thr Gln Asn Phe Glu Lys			
	355	360		365	
Phe	Glu Glu Ala Asp Asn Ser Ser Gln Ser Thr Ser Lys Ala Gly Pro				
	370	375		380	
Trp	Arg Lys Met Leu Ser Ser Lys Asp Leu Asn Phe Val Gly Tyr Thr				
	385	390		395	400
Tyr	Lys Asn Phe Glu Ile Val Asn Asp Tyr Gln Val Pro Gly Ile Ala				
	405	410		415	
Glu	Leu Lys Lys Lys Asp Thr Lys Pro Lys Arg Pro Ser Ile Lys Ser				
	420	425		430	
Leu	Phe Glu Asp Glu Ser Ser Asp Ser Ser Glu Ala Ala Thr Ser Gly				
	435	440		445	
Asp	Gln Ser Val Gln Gly Ser Phe Leu Asn Leu Leu Pro Pro Gln Leu				
	450	455		460	
Glu	Val Ser Gln Thr Gln Thr Glu Val Pro Pro Pro Lys Phe Thr				
	465	470		475	

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 500 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Glu Lys Val Lys Ala Ala Lys Lys Phe Ile Glu Asn His Tyr Arg
1	5 10 15
Ser	Gln Met Lys Asn Ile Gln Glu Arg Lys Glu Arg Arg Trp Val Leu
	20 25 30
Glu	Lys Gln Leu Ala Ser Ser Asp Val Pro Glu Glu Glu Gln Met Ser
	35 40 45
Leu	Ile Lys Asp Leu Glu Arg Lys Glu Thr Glu Phe Met Arg Leu Lys
	50 55 60
Arg	Asn Arg Ile Cys Val Asn Asp Phe Glu Leu Leu Thr Ile Ile Gly
	65 70 75 80
Arg	Gly Ala Tyr Gly Glu Val Gln Leu Cys Arg Glu Lys Lys Ser Glu
	85 90 95
Asn	Ile Tyr Ala Met Lys Lys Leu Lys Lys Ser Glu Met Leu Ser Arg
	100 105 110
Gly	Gln Val Glu His Val Arg Ala Glu Arg Asn Leu Leu Ala Glu Val
	115 120 125
Asp	Ser His Cys Ile Val Lys Leu Phe Tyr Ser Phe Gln Asp Ala Glu
	130 135 140

Tyr Leu Tyr Leu Ile Met Glu Tyr L u Pro Gly Gly Asp M t Met Thr  
 145 150 155 160  
 Leu Leu Met Arg Glu Asp Ile Leu Thr Glu Lys Val Ala Lys Phe Tyr  
 165 170 175  
 Ile Ala Gln Ser Val Leu Ala Ile Glu Ser Ile His Lys His Asn Tyr  
 180 185 190  
~~Ile His Arg Asp Ile Lys Pro Asp Asn Leu Leu Leu Asp Lys Asn Gly~~  
~~195 200 205~~  
 His Met Lys Leu Ser Asp Phe Gly Leu Cys Lys Pro Leu Asp Cys Ala  
 210 215 220  
 Thr Leu Ser Thr Ile Lys Glu Asn Glu Ser Met Asp Asp Val Ser Lys  
 225 230 235 240  
 Asn Ser Met Asp Ile Asp Ala Ser Leu Pro Asp Ala Gly Asn Gly His  
 245 250 255  
 Ser Trp Arg Ser Ala Arg Glu Gln Leu Gln His Trp Gln Arg Asn Arg  
 260 265 270  
 Arg Lys Leu Ala Phe Ser Thr Val Gly Thr Pro Asp Tyr Ile Ala Pro  
 275 280 285  
 Glu Val Leu Leu Lys Lys Gly Tyr Gly Met Glu Cys Asp Trp Trp Ser  
 290 295 300  
 Leu Gly Ala Ile Met Tyr Glu Met Leu Val Gly Tyr Pro Pro Phe Tyr  
 305 310 315 320  
 Ser Asp Asp Pro Ile Thr Thr Cys Arg Lys Ile Val His Trp Arg His  
 325 330 335  
 Tyr Leu Lys Phe Pro Asp Asp Ala Lys Leu Thr Phe Glu Ala Arg Asp  
 340 345 350  
 Leu Ile Cys Arg Leu Leu Cys Asp Val Glu His Arg Leu Gly Thr Gly  
 355 360 365  
 Gly Ala Glu Gln Ile Lys Val His Ala Trp Phe Lys Asp Val Glu Trp  
 370 375 380  
 Asp Arg Leu Tyr Glu Thr Asp Ala Ala Tyr Lys Pro Gln Val Asn Gly  
 385 390 395 400  
 Glu Leu Asp Thr Gln Asn Phe Met Lys Phe Asp Glu Ala Asn Pro Pro  
 405 410 415  
 Thr Pro Ser Arg Ser Gly Ser Gly Pro Ser Arg Lys Met Leu Thr Ser  
 420 425 430  
 Lys Asp Leu Ser Phe Val Gly Tyr Thr Tyr Lys Asn Phe Asp Ala Val  
 435 440 445  
 Lys Gly Leu Lys His Ser Phe Asp Arg Lys Gly Ser Thr Ser Pro Lys  
 450 455 460  
 Arg Pro Ser Leu Asp Ser Met Phe Asn Glu Asn Gly Met Asp Tyr Thr  
 465 470 475 480  
 Ala Lys His Ala Glu Glu Met Asp Val Gln Met Leu Thr Ala Asp Asp  
 485 490 495  
 Cys Met Ser Pro

500

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 564 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

Met Phe Ser Arg Ser Asp Arg Glu Val Asp Asp Leu Ala Gly Asn Met
1      5      10
Ser His Leu Gly Phe Tyr Asp Leu Asn Ile Pro Lys Pro Thr Ser Pro
20      25      30
Gln Ala Gln Tyr Arg Pro Ala Arg Lys Ser Glu Asn Gly Arg Leu Thr
35      40      45
Pro Gly Leu Pro Arg Ser Tyr Lys Pro Cys Asp Ser Asp Asp Gln Asp
50      55      60
Thr Phe Lys Asn Arg Ile Ser Leu Asn His Ser Pro Lys Lys Leu Pro
65      70      75      80
Lys Asp Phe His Glu Arg Ala Ser Gln Ser Lys Thr Gln Arg Val Val
85      90      95
Asn Val Cys Gln Leu Tyr Phe Leu Asp Tyr Tyr Cys Asp Met Phe Asp
100     105     110
Tyr Val Ile Ser Arg Arg Gln Arg Thr Lys Gln Val Leu Arg Tyr Leu
115     120     125
Glu Gln Gln Arg Ser Val Lys Asn Val Ser Asn Lys Val Leu Asn Glu
130     135     140
Glu Trp Ala Leu Tyr Leu Gln Arg Glu His Glu Val Leu Arg Lys Arg
145     150     155     160
Arg Leu Lys Pro Lys His Lys Asp Phe Gln Ile Leu Thr Gln Val Gly
165     170     175
Gln Gly Gly Tyr Gly Gln Val Tyr Leu Ala Lys Lys Lys Asp Ser Asp
180     185     190
Glu Ile Cys Ala Leu Lys Ile Leu Asn Lys Lys Leu Leu Phe Lys Leu
195     200     205
Asn Glu Thr Asn His Val Leu Thr Glu Arg Asp Ile Leu Thr Thr Thr
210     215     220
Arg Ser Asp Trp Leu Val Lys Leu Leu Tyr Ala Phe Gln Asp Pro Glu
225     230     235     240
Ser Leu Tyr Leu Ala Met Glu Phe Val Pro Gly Gly Asp Phe Arg Thr
245     250     255
L u Leu Ile Asn Thr Arg Ile Leu Lys Ser Gly His Ala Arg Ph Tyr
260     265     270

```

Il Ser Glu Met Phe Cys Ala Val Asn Ala Leu His Glu Leu Gly Tyr  
 275 280 285  
 Thr His Arg Asp Leu Lys Pro Glu Asn Ph L u Ile Asp Ala Thr Gly  
 290 295 300  
 His Ile Lys Leu Thr Asp Phe Gly Leu Ala Ala Gly Thr Val S r Asn  
 305 310 315 320  
 Glu Arg Ile Glu Ser Met Lys Ile Arg Leu Glu Glu Val Lys Asn Leu  
 325 330 335  
 Gln Phe Pro Ala Phe Thr Glu Arg Ser Ile Glu Asp Arg Ser Lys Ile  
 340 345 350  
 Tyr His Asn Met Arg Lys Thr Glu Ile Asn Tyr Ala Asn Ser Met Val  
 355 360 365  
 Gly Ser Pro Asp Tyr Met Ala Leu Glu Val Leu Glu Gly Lys Lys Tyr  
 370 375 380  
 Asp Phe Thr Val Asp Tyr Trp Ser Leu Gly Cys Met Leu Phe Glu Ser  
 385 390 395 400  
 Leu Val Gly Tyr Thr Pro Phe Ser Gly Ser Ser Thr Asn Glu Thr Tyr  
 405 410 415  
 Glu Asn Leu Arg Tyr Trp Lys Lys Thr Leu Arg Arg Pro Arg Thr Glu  
 420 425 430  
 Asp Arg Arg Ala Ala Phe Ser Asp Arg Thr Trp Asp Leu Ile Thr Arg  
 435 440 445  
 Leu Ile Ala Asp Pro Ile Asn Arg Val Arg Ser Phe Glu Gln Val Arg  
 450 455 460  
 Lys Met Ser Tyr Phe Ala Glu Ile Asn Phe Glu Thr Leu Arg Thr Ser  
 465 470 475 480  
 Ser Pro Pro Phe Ile Pro Gln Leu Asp Asp Glu Thr Asp Ala Gly Tyr  
 485 490 495  
 Phe Asp Asp Phe Thr Asn Glu Glu Asp Met Ala Lys Tyr Ala Asp Val  
 500 505 510  
 Phe Lys Arg Gln Asn Lys Leu Ser Ala Met Val Asp Asp Ser Ala Val  
 515 520 525  
 Asp Ser Lys Leu Val Gly Phe Thr Phe Arg His Arg Asp Gly Lys Gln  
 530 535 540  
 Gly Ser Ser Gly Ile Leu Tyr Asn Gly Ser Glu His Ser Asp Pro Phe  
 545 550 555 560  
 Ser Thr Phe Tyr

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 561 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID N :16:

Met Ala Gly Asn Met Ser Asn Leu Ser Phe Asp Gly His Gly Thr Pro  
 1 5 10 15  
 Gly Gly Thr Gly Leu Phe Pro Asn Gln Asn Il Thr Lys Arg Arg Thr  
 20 25 30  
 Arg Pro Ala Gly Ile Asn Asp Ser Pro Ser Pro Val Lys Pro Ser Phe  
 35 40 45  
 Phe Pro Tyr Glu Asp Thr Ser Asn Met Asp Ile Asp Glu Val Ser Gln  
 50 55 60  
 Pro Asp Met Asp Val Ser Asn Ser Pro Lys Lys Leu Pro Pro Lys Phe  
 65 70 75 80  
 Tyr Glu Arg Ala Thr Ser Asn Lys Thr Gln Arg Val Val Ser Val Cys  
 85 90 95  
 Lys Met Tyr Phe Leu Glu Tyr Tyr Cys Asp Met Phe Asp Tyr Val Ile  
 100 105 110  
 Ser Arg Arg Gln Arg Thr Lys Gln Val Leu Glu Tyr Leu Gln Gln Gln  
 115 120 125  
 Ser Gln Leu Pro Asn Ser Asp Gln Ile Lys Leu Asn Glu Glu Trp Ser  
 130 135 140  
 Ser Tyr Leu Gln Arg Glu His Gln Val Leu Arg Lys Arg Arg Leu Lys  
 145 150 155 160  
 Pro Lys Asn Arg Asp Phe Glu Met Ile Thr Gln Val Gly Gln Gly Gly  
 165 170 175  
 Tyr Gly Gln Val Tyr Leu Ala Arg Lys Lys Asp Thr Lys Glu Val Cys  
 180 185 190  
 Ala Leu Lys Ile Leu Asn Lys Lys Leu Leu Phe Lys Leu Asn Glu Thr  
 195 200 205  
 Lys His Val Leu Thr Glu Arg Asp Ile Leu Thr Thr Thr Arg Ser Glu  
 210 215 220  
 Trp Leu Val Lys Leu Leu Tyr Ala Phe Gln Glu Leu Gln Ser Leu Tyr  
 225 230 235 240  
 Leu Ala Met Glu Phe Val Pro Gly Gly Asp Phe Arg Thr Leu Leu Ile  
 245 250 255  
 Asn Thr Arg Cys Leu Lys Ser Gly His Ala Arg Phe Tyr Ile Ser Glu  
 260 265 270  
 Met Phe Cys Ala Val Asn Ala Leu His Asp Leu Gly Tyr Thr His Arg  
 275 280 285  
 Asp Leu Lys Pro Glu Asn Phe Leu Ile Asp Ala Lys Gly His Ile Lys  
 290 295 300  
 Leu Thr Asp Phe Gly Leu Ala Ala Gly Thr Ile Ser Asn Glu Arg Ile  
 305 310 315 320  
 Glu Ser Met Lys Ile Arg Leu Glu Lys Ile Lys Asp Leu Glu Phe Pro  
 325 330 335  
 Ala Phe Thr Glu Lys Ser Ile Glu Asp Arg Arg Lys Met Tyr Asn Gln

340							345						350					
Leu Arg Glu Lys Glu Ile Asn Tyr Ala Asn Ser Met Val Gly Ser Pro 355 360 365																		
Asp Tyr M t Ala Leu Glu Val Leu Glu Gly Lys Lys Tyr Asp Phe Thr 370 375 380																		
Val Asp Tyr Trp Ser Leu Gly Cys Met Leu Phe Glu Ser Leu Val Gly 385 390 395 400																		
Tyr Thr Pro Phe Ser Gly Ser Ser Thr Asn Glu Thr Tyr Asp Asn Leu 405 410 415																		
Arg Arg Trp Lys Gln Thr Leu Arg Arg Pro Arg Gln Ser Asp Gly Arg 420 425 430																		
Ala Ala Phe Ser Asp Arg Thr Trp Asp Leu Ile Thr Arg Leu Ile Ala 435 440 445																		
Asp Pro Ile Asn Arg Leu Arg Ser Phe Glu His Val Lys Arg Met Ser 450 455 460																		
Tyr Phe Ala Asp Ile Asn Phe Ser Thr Leu Arg Ser Met Ile Pro Pro 465 470 475 480																		
Phe Thr Pro Gln Leu Asp Ser Glu Thr Asp Ala Gly Tyr Phe Asp Asp 485 490 495																		
Phe Thr Ser Glu Ala Asp Met Ala Lys Tyr Ala Asp Val Phe Lys Arg 500 505 510																		
Gln Asp Lys Leu Thr Ala Met Val Asp Asp Ser Ala Val Ser Ser Lys 515 520 525																		
Leu Val Gly Phe Thr Phe Arg His Arg Asn Gly Lys Gln Gly Ser Ser 530 535 540																		
Gly Ile Leu Phe Asn Gly Leu Glu His Ser Asp Pro Phe Ser Thr Phe 545 550 555 560																		
Tyr																		



International Application No: PCT/ /

**MICROORGANISMS**Optional Sheet in connection with the microorganism referred to on page 108, lines 1-20 of the description \***A. IDENTIFICATION OF DEPOSIT \***

Further deposits are identified on an additional sheet \*

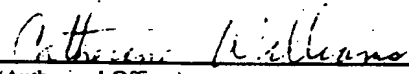
Name of depositary institution \*

American Type Culture Collection

Address of depositary institution (including postal code and country) \*

12301 Parklawn Drive  
Rockville, MD 20852  
USDate of deposit \* March 24, 1995 Accession Number \* 69769**B. ADDITIONAL INDICATIONS** \* (leave blank if not applicable). This information is continued on a separate attached sheet**C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE \*** (If the indications are on all designated States)**D. SEPARATE FURNISHING OF INDICATIONS \*** (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later \* (Specify the general nature of the indications e.g., "Accession Number of Deposit")

**E.** ☒ This sheet was received with the International application when filed (to be checked by the receiving Office)  
(Authorized Officer)☐ The date of receipt (from the applicant) by the International Bureau \*

was

\_\_\_\_\_  
(Authorized Officer)

Form PCT/RO/134 (January 1981)

WHAT IS CLAIMED IS:

1. A purified lats protein.
- 5           2. The protein of claim 1 which is a human protein.
3. The protein of claim 1 which is a *D. melanogaster* protein.
- 10           4. The protein of claim 1 which is a mouse protein.
5. The protein of claim 1 which is a mammalian
- 15 protein.
6. The protein of claim 2 which comprises the amino acid sequence substantially as set forth in SEQ ID NO:4.
- 20           7. A purified protein encoded by a nucleic acid hybridizable to the lats DNA sequence in plasmid PBS(KS)-h-lats as deposited with the ATCC and assigned accession number 69769.
- 25           8. A purified protein encoded by a nucleic acid hybridizable to a DNA having a sequence consisting of the coding region of SEQ ID NO:7.
9. The protein of claim 2 which is encoded by plasmid pBS(KS)-h-lats as deposited with the ATCC and assigned accession number 69769.
- 30           10. A purified derivative or analog of the protein
- 35 of claim 1, which displays one or more functional activities of a lats protein.

11. The derivative or analog of claim 10 which is able to be bound by an antibody directed against a lats protein.

5 12. A purified fragment of a lats protein comprising a domain of the lats protein selected from the group consisting of a lats C-terminal domain 3 (LCD3), lats C-terminal domain 2 (LCD2), lats C-terminal domain 1 (LCD1), kinase domain, a kinase subdomain, lats flanking domain  
10 (LFD), lats split domain 1 (LSD1), lats split domain 2 (LSD2), SH3-binding domain, and opa repeat domain.

13. A molecule comprising the fragment of claim 12.

15

14. A protein comprising an amino acid sequence that has at least 60% identity to a domain of a lats protein, in which the percentage identity is determined over an amino acid sequence of identical size to the domain.

20

15. A protein comprising an amino acid sequence that has at least 90% identity to a domain of a lats protein, in which the percentage identity is determined over an amino acid sequence of identical size to the domain.

25

16. The derivative or analog of claim 10, which inhibits proliferation of a cell.

17. A chimeric protein comprising a fragment of a  
30 lats protein consisting of at least 6 amino acids fused via a covalent bond to an amino acid sequence of a second protein, in which the second protein is not a lats protein.

18. The chimeric protein of claim 17 in which the  
35 fragment of a lats protein is a fragment capable of being bound by an anti-lats antibody.

19. The fragment of claim 12 which additionally lacks one or more domains of the lats protein.

20. An antibody which is capable of binding a lats  
5 protein.

21. The antibody of claim 20 which is monoclonal.

22. A molecule comprising a fragment of the  
10 antibody of claim 21, which fragment is capable of binding a lats protein.

23. An isolated nucleic acid comprising a nucleotide sequence encoding a lats protein.

15

24. The nucleic acid of claim 23 which is a DNA.

25. An isolated nucleic acid comprising a nucleotide sequence complementary to the nucleotide sequence  
20 of claim 23.

26. The nucleic acid of claim 23 in which the lats protein is a human lats protein.

25 27. An isolated nucleic acid comprising the lats coding sequence contained in plasmid pBS(KS)-h-lats as deposited with the ATCC and assigned accession number 69769.

28. An isolated nucleic acid hybridizable to the  
30 lats DNA sequence in plasmid pBS(KS)-h-lats as deposited with the ATCC and assigned accession number 69769.

29. An isolated nucleic acid hybridizable to a DNA having a sequence consisting of the coding region of  
35 SEQ ID NO:7.

30. An isolated nucleic acid comprising a fragment of a lats g ne consisting of at least 8 nucleotid s.

31. An is lated nucleic acid comprising a  
5 nucleotide sequence encoding a fragment of a lats protein that displays one or more functional activities of the lats protein.

32. An isolated nucleic acid comprising a  
10 nucleotide sequence encoding the chimeric protein of claim 17.

33. An isolated nucleic acid comprising a nucleotide sequence encoding a protein, said protein  
15 comprising the amino acid sequence of SEQ ID NO:4.

34. An isolated nucleic acid comprising a nucleotide sequence encoding the fragment of claim 12.

20 35. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 14.

36. A recombinant cell containing the nucleic acid of claim 23, in which the nucleotide sequence encoding the  
25 lats protein is under the control of a promoter that is not a native lats gene promoter.

37. A recombinant cell containing a nucleic acid vector comprising the nucleic acid of claim 26.  
30

38. A recombinant cell containing the nucleic acid of claim 34.

39. A recombinant cell containing the nucleic acid  
35 of claim 35.

40. A method of producing a lats protein comprising growing a recombinant cell containing the nucleic acid of claim 23, in which the nucleotide sequence encoding the lats protein is under the control of a promoter that is not a native lats gene promoter, such that the encoded lats protein is expressed by the cell, and recovering the expressed lats protein.

41. A method of producing a lats protein comprising growing a recombinant cell containing a nucleic acid vector comprising the nucleic acid of claim 26 such that the encoded lats protein is expressed by the cell, and recovering the expressed lats protein.

42. A method of producing a lats fragment comprising growing a recombinant cell containing the nucleic acid of claim 34 such that the encoded lats fragment is expressed by the cell, and recovering the expressed lats fragment.

43. A method of producing a protein comprising a fragment of a lats protein, which method comprises growing a recombinant cell containing the nucleic acid of claim 35 such that the encoded protein is expressed by the cell, and recovering the expressed protein.

44. The product of the process of claim 40.

45. The product of the process of claim 41.

46. The product of the process of claim 42.

47. The product of the process of claim 43.

48. A pharmaceutical composition comprising a therapeutically effective amount of a lats protein; and a pharmaceutically acceptable carrier.

49. The composition of claim 48 in which the lats protein is a human lats protein.

50. A pharmaceutical composition comprising a  
5 therapeutically effective amount of the fragment of claim 12;  
and a pharmaceutically acceptable carrier.

51. A pharmaceutical composition comprising a  
therapeutically effective amount of the protein of claim 14;  
10 and a pharmaceutically acceptable carrier.

52. A pharmaceutical composition comprising a  
therapeutically effective amount of the chimeric protein of  
claim 17; and a pharmaceutically acceptable carrier.

15

53. A pharmaceutical composition comprising a  
therapeutically effective amount of the nucleic acid of claim  
23; and a pharmaceutically acceptable carrier.

20

54. A pharmaceutical composition comprising a  
therapeutically effective amount of the nucleic acid of claim  
35; and a pharmaceutically acceptable carrier.

55. A pharmaceutical composition comprising a  
25 therapeutically effective amount of the recombinant cell of  
claim 36; and a pharmaceutically acceptable carrier.

56. A pharmaceutical composition comprising a  
therapeutically effective amount of an antibody that  
30 immunospecifically binds to a lats protein; and a  
pharmaceutically acceptable carrier.

57. A pharmaceutical composition comprising a  
therapeutically effective amount of a fragment or derivative  
35 of an antibody that immunospecifically binds to a lats  
protein, said fragment or derivative containing the binding

domain of the antibody; and a pharmaceutically acceptable carrier.

58. A method of treating or preventing a disease or disorder involving cell overproliferation in a subject comprising administering to a subject in which such treatment or prevention is desired a therapeutically effective amount of a molecule that promotes lats function.

59. The method according to claim 58 in which the disease or disorder is a malignancy.

60. The method according to claim 59 in which the disease or disorder is selected from the group consisting of bladder cancer, breast cancer, colon cancer, leukemia, lung cancer, melanoma, pancreatic cancer, sarcoma, and uterine cancer.

61. The method according to claim 58 in which the subject is a human.

62. The method according to claim 58 in which the disease or disorder is selected from the group consisting of premalignant conditions, benign tumors, hyperproliferative disorders, and benign dysproliferative disorders.

63. The method according to claim 58 in which the molecule that promotes lats function is selected from the group consisting of a lats protein, a lats derivative or analog that is active in inhibiting cell proliferation, a nucleic acid encoding a lats protein, and a nucleic acid encoding a lats derivative or analog that is active in inhibiting cell proliferation.

64. The method according to claim 58 in which the molecule that promotes lats function is a lats derivative or



anal g that comprises a kinase domain of a lats protein that has b n mutated so as to be dominantly active.

65. Th method according to claim 58 in which the  
5 molecule that promotes lats function is the protein of claim 14.

66. A method of treating or preventing a disease or disorder involving a deficiency in cell proliferation or  
10 in which cell proliferation is desirable for treatment or prevention in a subject comprising administering to a subject in which such treatment or prevention is desired a therapeutically effective amount of a molecule that inhibits lats function.

15

67. The method according to claim 66 in which the molecule that inhibits lats function is selected from the group consisting of an anti-lats antibody or a fragment or derivative thereof containing the binding region thereof, a  
20 lats derivative or analog that is capable of being bound by an anti-lats antibody and that is a dominant-negative protein kinase, a lats antisense nucleic acid, and a nucleic acid comprising at least a portion of a lats gene into which a heterologous nucleotide sequence has been inserted such that  
25 said heterologous sequence inactivates the biological activity of the at least a portion of the lats gene, in which the lats gene portion flanks the heterologous sequence so as to promote homologous recombination with a genomic lats gene.

30 68. The method according to claim 66 in which the molecule that inhibits lats function is an oligonucleotide which (a) consists of at least six nucleotides; (b) comprises a sequence complementary to at least a portion of an RNA transcript of a lats gene; and (c) is hybridizable to the RNA  
35 transcript under moderately stringent conditions.

69. The method according to claim 66 in which the disease or disorder is selected from the group consisting of degenerative disorders, growth deficiencies, hypoproliferative disorders, physical trauma, lesions, and wounds.

70. An isolated oligonucleotide consisting of at least six nucleotides, and comprising a sequence complementary to at least a portion of an RNA transcript of a lats gene, which oligonucleotide is hybridizable to the RNA transcript under moderately stringent conditions.

71. A pharmaceutical composition comprising the oligonucleotide of claim 70; and a pharmaceutically acceptable carrier.

72. A method of inhibiting the expression of a nucleic acid sequence encoding a lats protein in a cell comprising providing the cell with an effective amount of the oligonucleotide of claim 70.

73. A method of diagnosing a disease or disorder characterized by an aberrant level of lats RNA or protein in a subject, comprising measuring the level of lats RNA or protein in a sample derived from the subject, in which an increase or decrease in the level of lats RNA or protein, relative to the level of lats RNA or protein found in an analogous sample not having the disease or disorder indicates the presence of the disease or disorder in the subject.

74. A method of diagnosing or screening for the presence of or a predisposition for developing a disease or disorder involving cell overproliferation in a subject comprising measuring the level of lats protein, lats RNA or lats functional activity in a sample derived from the subject, in which a decrease in the level of lats protein,

lats RNA, or lats functional activity in the sample, relative to the level of lats protein, lats RNA, or lats functional activity found in an analogous sample not having the disease or disorder or a predisposition for developing the disease or disorder, indicates the presence of the disease or disorder or a predisposition for developing the disease or disorder.

75. A method of diagnosing or screening for the presence of or a predisposition for developing a disease or disorder involving cell overproliferation in a subject comprising detecting one or more mutations in lats DNA, RNA or protein derived from the subject in which the presence of said one or more mutations indicates the presence of the disease or disorder or a predisposition for developing the disease or disorder.

76. A method of diagnosing or screening for the presence of or a predisposition for developing a disease or disorder involving a deficiency in cell proliferation or in which cell proliferation is desirable for treatment or prevention in a subject comprising measuring the level of lats protein, lats RNA or lats functional activity in a sample derived from the subject, in which an increase in the level of lats protein, lats RNA, or lats functional activity in the sample, relative to the level of lats protein, lats RNA, or lats functional activity found in an analogous sample not having the disease or disorder or a predisposition for developing the disease or disorder, indicates the presence of the disease or disorder or a predisposition for developing the disease or disorder.

77. A kit comprising in one or more containers a molecule selected from the group consisting of an anti-lats antibody, a nucleic acid probe capable of hybridizing to a lats RNA, or a pair of nucleic acid primers capable of priming amplification of at least a portion of a lats nucleic acid.

78. A kit comprising in a container a  
th rapeutically effective amount of a lats protein.

79. A m thod f increasing cell growth in animals  
5 or plants comprising inhibiting lats expression or activity  
in said animals or plants.

80. The method of claim 79 in which cell growth is  
increased in an edible plant.

10

81. The method of claim 79 in which cell growth is  
increased in a farm animal.

82. A method of identifying a molecule that  
15 specifically binds to a ligand selected from the group  
consisting of a lats protein, a fragment of a lats protein  
comprising a domain of the protein, and a nucleic acid  
encoding the protein or fragment, comprising

20

(a) contacting said ligand with a plurality of  
molecules under conditions conducive to  
binding between said ligand and the molecules;  
and

(b) identifying a molecule within said plurality  
that specifically binds to said ligand.

25

83. A recombinant non-human animal or plant that  
is the product of a process comprising introducing a nucleic  
acid encoding at least a domain of a lats protein into the  
plant or animal.

30

84. A recombinant plant containing and capable of  
expressing a lats antisense nucleic acid.

85. A recombinant non-human animal or plant in  
35 which a lats gene has been inactivated by a method comprising  
introducing a nucleic acid into the plant or animal or an  
ancestor thereof, which nucleic acid comprises a non-lats

s quence flanked by lats genomic sequences that promote homologous r combination.

86. A method of identifying a tumor suppressor  
5 gene comprising (a) identifying an overproliferation phenotype in a genetic mosaic; and (b) isolating a gene that is mutated in cells exhibiting said overproliferation phenotype.

10 87. The method of claim 86 in which the genetic mosaic is an animal containing (a) a nucleic acid encoding and capable of expressing a recombinase, and (b) intrachromosomal insertions of a target site at which the recombinase promotes recombination, on the homologous arms of  
15 both of a set of parental chromosomes; and the genetic mosaic has been produced by a method comprising inducing expression of the recombinase.

88. The method of claim 87 in which the  
20 recombinase is an FLP recombinase, and the target site is an FRT site.

89. The method according to claim 87 in which the recombinase is a Cre recombinase, and the target site is a  
25 lox site.

90. The method of claim 86 in which the overproliferation phenotype is the formation of overproliferated outgrowth tissue.

30

91. The method of claim 86 in which the overproliferation phenotype is the formation of a normal structure of larger than normal size.

35 92. A non-human mammal comprising (a) a nucleic acid sequence encoding a recombinase operably linked to a promoter; and (b) intrachromosomal insertions into the

homologous arms of both of a set of parental chromosomes, of a target site at which the recombinase can promote recombination.

5                   93. The mammal of claim 92 which is heterozygous for an induced mutation.

                  94. The mammal of claim 93 in which the sequence encoding the recombinase is operably linked to an inducible  
10 promoter.

                  95. A method of making a genetic mosaic comprising inducing expression of the recombinase in the mammal of claim 93.

15

                  96. A method for identifying a gene with an identifiable mutant phenotype comprising:

- 20                   (a) identifying a mutant phenotype in a genetic mosaic animal, said genetic mosaic animal having been produced by a method comprising recombinantly expressing a recombinase within a cell of the animal so as to promote recombination at intrachromosomally inserted target sites on the homologous arms of both of  
25 a set of parental chromosomes; and  
                  (b) isolating a gene that is mutated in cells exhibiting said mutant phenotype.

30                   97. A method for identifying a gene with an identifiable mutant phenotype comprising:

- 35                   (a) identifying a mutant phenotype in a cultured cell, said cultured cell having been produced by a method comprising recombinantly expressing a recombinase within said cultured cell so as to promote recombination at intrachromosomally inserted target sites on

the homologous arms of both of a set of parental chromosomes; and

- (b) isolating a gene that is mutated in cells exhibiting said mutant phenotype.

5

98. The method of claim 97 in which the mutant phenotype is a transformed phenotype.

99. The mammal of claim 92 in which the promoter  
10 is not a native recombinase gene promoter.

100. A method of inhibiting cellular senescence in a subject comprising administering to a subject in which such inhibition is desired an amount of a molecule that inhibits  
15 lats function, effective to inhibit cellular senescence.

101. A method of inhibiting cellular senescence in cells *in vitro* comprising contacting cells *in vitro* with an amount of a molecule that inhibits lats function, effective  
20 to inhibit cellular senescence.

102. The method according to claim 100 in which the molecule that inhibits lats function is selected from the group consisting of an anti-lats antibody or a fragment or  
25 derivative thereof containing the binding region thereof, a lats derivative or analog that is capable of being bound by an anti-lats antibody and that is a dominant-negative protein kinase, a lats antisense nucleic acid, and a nucleic acid comprising at least a portion of a lats gene into which a  
30 heterologous nucleotide sequence has been inserted such that said heterologous sequence inactivates the biological activity of the at least a portion of the lats gene, in which the lats gene portion flanks the heterologous sequence so as to promote homologous recombination with a genomic lats gene.

35

103. The method according to claim 101 in which the molecule that inhibits lats function is selected from the

gr up consisting of an anti-lats antibody or a fragment or derivative thereof containing the binding region thereof, a lats derivative or analog that is capable of being bound by an anti-lats antibody and that is a dominant-negative protein kinase, a lats antisense nucleic acid, and a nucleic acid comprising at least a portion of a lats gene into which a heterologous nucleotide sequence has been inserted such that said heterologous sequence inactivates the biological activity of the at least a portion of the lats gene, in which the lats gene portion flanks the heterologous sequence so as to promote homologous recombination with a genomic lats gene.

15

BAD ORIGINAL

20

25

30

35



1/43

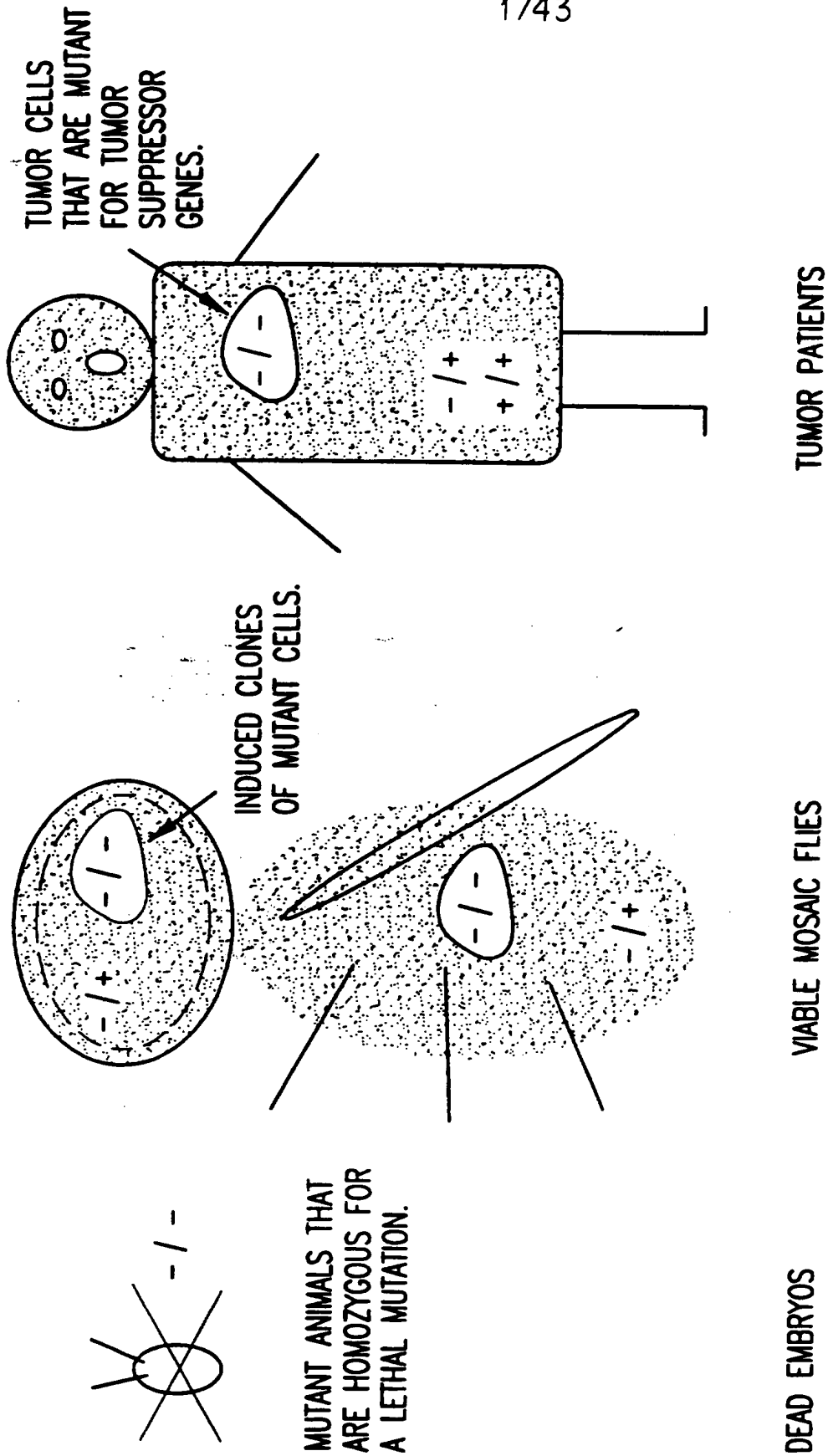


FIG.1A

2/43

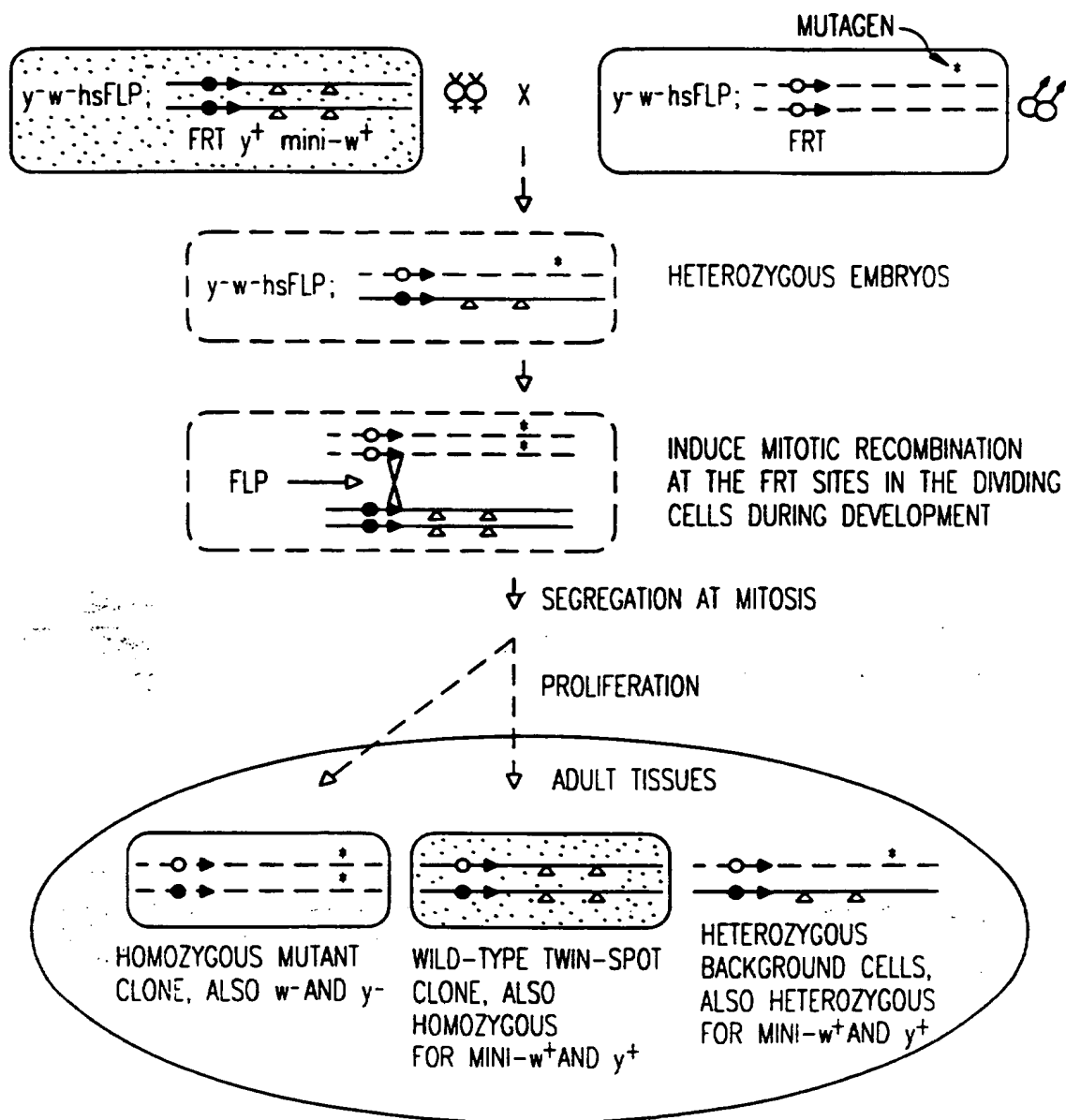


FIG.1B

3/43



FIG. 2C



FIG. 2B



FIG. 2A

4/43



FIG. 2F



FIG. 2E

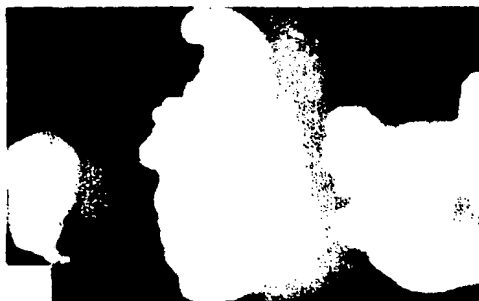


FIG. 2D

5/43



FIG. 2I

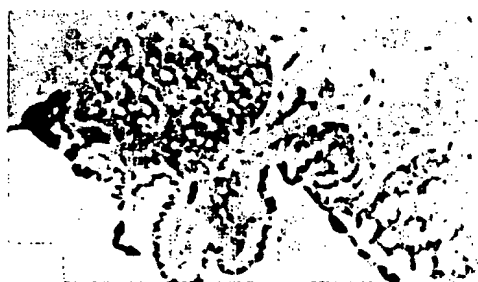


FIG. 2H

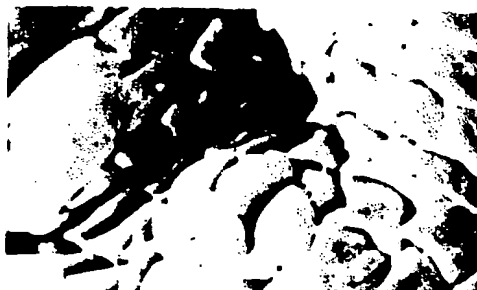


FIG. 2G

6/43

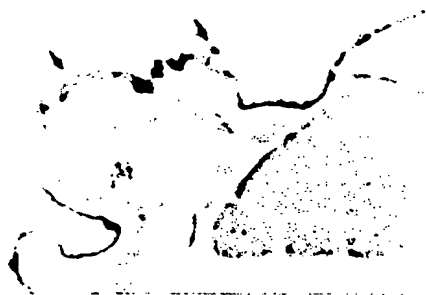


FIG. 2L



FIG. 2K

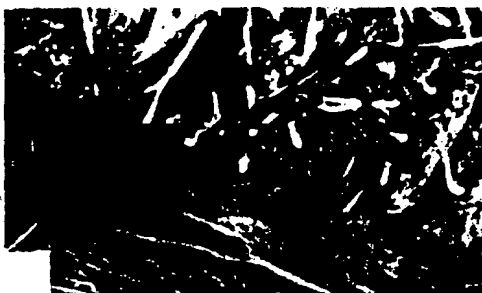


FIG. 2J

7/43

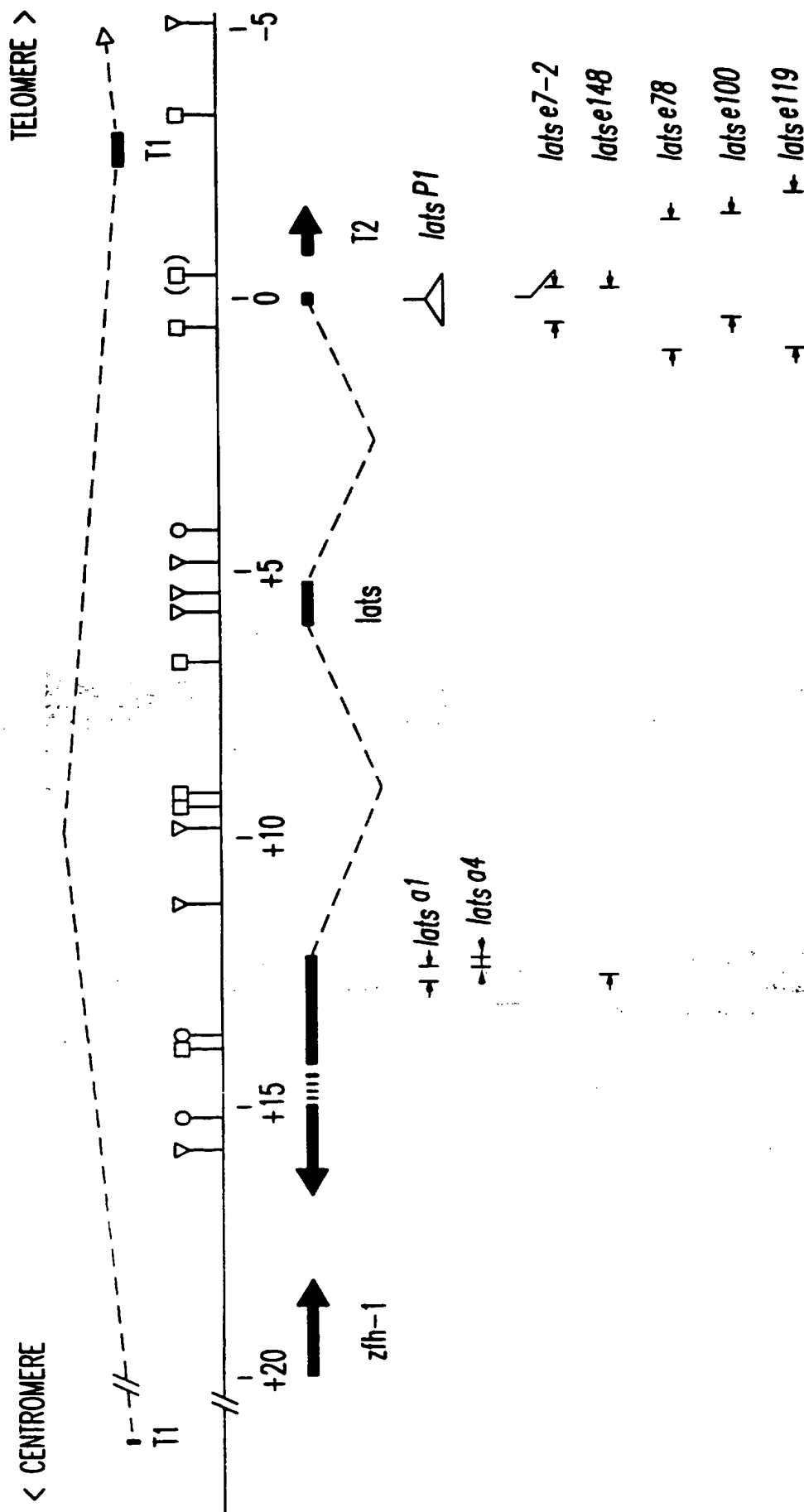


FIG.3

8/43

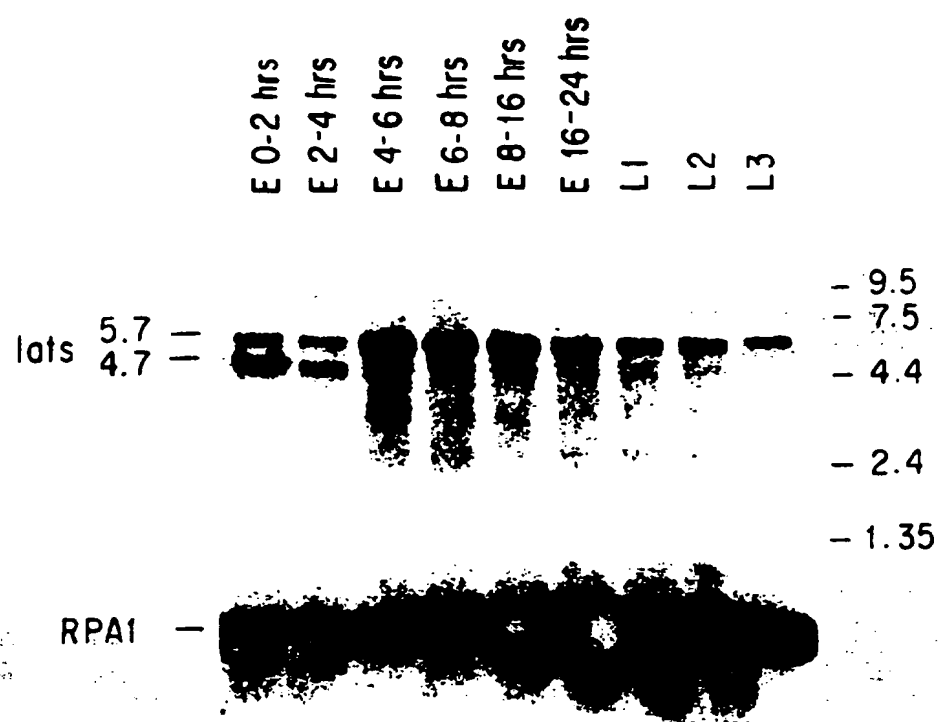


FIG.4



[illegible]

FIG. 5A

**SUBSTITUTE SHEET (RULE 26)**

## BAD ORIGINAL

10/43

G

1651 TTCTCCGTGCGCAAGCGGTTTTCAGTGAGGTGGCTCCACCGGCGCGCGGCCACGCAATCCCACCGCCTCCAGCGC  
184 S P S P S G F S E V A P P A P P P R N P T A S S A

1726 GGCCACGCCCCACCGCCAGTGGCGCCACCAGCCAGGCGTACGTGAAGCGGCGATCACCGGCCCTGAACAACCG  
209 A T P P P P V P P T S Q A Y V K R R S P A L N N R

lat-al deletion

1801 CCCGCCGGGATAGCGCCACCCCACTCAGCGAGGCAACTCACCTGTSATAACCCAAAACGGGCTGAAGAACCCGC  
234 P P A I A P P T Q R G N S P V I T Q N G L K N P Q

1876 GCAGCAGTTGACGCAGCAGCTGAAGTCCCTGAACCTATACCCAGCGGAGGCAGTGGAGCAGTGGTGGAGCCACC  
259 Q Q L T Q Q L K S L N L Y P G G G S G A V V E P P

1951 GCGGCCCTACCTAATTCAAGCGGAGCGGAGGAGCAGCACCGCGCGGCCACCCAGTTACAGGCCTCCAT  
284 P P Y L I Q G G A G G A A P P P P P S Y T A S M

2026 GCAGTCGCGGCAGTCGCCACACAATCCCAACAATCGGACTACAGGAAATCCCCGAGCAGTGGGATATACTCGGC  
309 Q S R Q S P T Q S Q Q S D Y R K S P S S G I Y S A

2101 CACCTCGCGGGCTCGCCGAGCCCCATAACTGTGTGCTGCCGCGGCGCGGCTGGCGAAGCCACAACCACGAGT  
334 T S A G S P S P I T V S L P P A P L A K P Q P R V

2176 CTACCAGGCCAGGAGTCAGCAGCCGATCATTCAGAGTGTGAAGAGCAGCAGGTCCAAAAGCCCGTGTGCA  
359 Y Q A R S Q Q P I I M Q S V K S T Q V Q K P V L Q

2251 AACAGCAGTGGCGCCCAATCGCCATCGAGTGCTCGGCCAGCAATTCACCAGTCCACGTGCTGGCCGCTCCACC  
384 T A V A P Q S P S S A S A S N S P V H V L A A P P

2306 CTCTTACCCTCAGAAGTCCGCGGCAGTGGTGCAGCAGCAGCAACAGGCAGCAGCGCGGCCACCAGCAGCAGCA  
409 S Y P Q K S A A V V Q Q Q Q Q A A A A H Q Q Q H

2401 TCAGCACCAGCAATCCAAACCACCAACGCCAACCACACCGCCCTTGGTGGGTCTGAACAGCAAGCCCAATTGCCT  
434 Q H Q Q S K P P T P T P P L V G L N S K P N C L

2476 GGAGCCACCGTCCTATGCCAAGAGCATGCAGGCCAAGCGGCCACGGTGGTACAGCAGCAGCAACAGCAGCAGCA  
459 E P P S Y A K S M Q A K A A T V V Q Q Q Q Q Q Q Q

AAC

G G A

G

2551 ACAACAGCAGGTCCAGCAGCAGCAGGTGCAACAGCAGCAGCAACAGCAGCAACAGCAACTGCAGGCCTTGAGGGT  
484 Q Q Q V Q Q Q Q V Q Q Q Q Q Q Q Q L Q A L R V

GGGAGCGGGATCAAC

2626 GCTCCAGGCACAGGCTCAGAGGGAGCGGGATCAACGGGAGCGGGAACGGGATCAGCAGAAGCTGGCCAAACGGAAA  
559 L Q A Q A Q R E R D Q R E R E R D Q Q K L A N G N

2701 TCCTGGCCGCGCAGATGCTTCGCGCGCGCCCTATCAGAGCAACAACAACAACAGCGAGATCAAACCGCGAG  
534 P G R Q M L P P P P Y Q S N N N N N S E I K F F S

2776 CTGCAACAACAACAACATACAGATAAGCAACAGCAACCTGGCGACGACACCACCCATTCCGCCTGCCAAATACAA  
559 C N N N N I Q I S N S N L A T T P P I P P A K Y N

2851 TAACAACTCCTCCAACACGGGCGGAATAGCTCGGGCGGCAGCAACGGATCCACCGGCACCACCGCCTCCTCGTC  
584 N N S S N T G A N S S G G S N G S T G T T A S S S

2926 GACCAGCTGCAAGAAGATCAAGCAGCGCTCGCCCATCCCGGAGCGCAAGAAGATCTCCAAGGAGAAGGAGGAGGA  
609 T S C K K I K H A S P I P E R K K I S K E K E E E

FIG.5B

SUBSTITUTE SHEET (RULE 26)

BAD ORIGINAL

11/43

3001 GCGCAAGGAGTTCGCGATCAGGCAGTACTCGCCGCAAGCCTTCAAGTTCTTCATGGAGCAGCACATAGAGAACGT  
634 R K E F R I R Q Y S P Q A F K F F M E Q H I E N V

3076 GATCAAGTCGTATCGCCAGCGCACGTATCGCAAGAATCAGCTGGAGAAGGAGATGCACAAAGTGGGACTGCCCGA  
709 I K S Y R Q R T Y R K N Q L E K E M H K V G L P D

3151 TCAGACCCAAATCGAGATGAGGAAAATGCTGAACCAAAAGGAGAGCAACTACATTGATTGAAGCGCGCCAAGAT  
684 Q T Q I E M R K M L N Q K E S N Y I R L K R A K M

3226 GGACAAGAGCATGTTCTGTCAAACTGAAGCCCAATTGGAGTGGGTGCATTGGCGAGGTAACGCTGGTGAGCAAT  
759 D K S M F V K L K P I G V G A F G E V T L V S K I

3301 CGATACCTCGAACCATTGTATGCGATGAAAACCTCGCGAAAGCGGACGTTCTCAAGCGGAATCAGGTGCCACA  
734 D T S N H L Y A M K T L R K A D V L K R N Q V A H

3376 CGTGAAGGCCGAGAGGGATATCCTCGCGAAGCCGACAATAACTGGGTGGTGAAGTTGTTACTACAGCTTCCAGGA  
809 V K A E R D I L A E A D N N W V V K L Y Y S F Q D  
Intron 3

3451 CAAGGATAATCTGTACTTTGTGATGGACTACATACCGTGGTGGTGTCTGTCTGCTCATCAAACTGGGCAT  
784 K D N L Y F V M D Y I P G G D L M S L L I K L G I

3526 TTTGAGGAGGAAGTGGCCAGATTCTACATCGCCGAGGTACCTGCGCCGTGGACAGCGTTCACAAAATFFFCTT  
809F F E E E L A R E Y I A E V T C A V D S V H K M G F  
Intron 4

3601 CATTACAGAGACATCAAGCCTGACAACATACTCATCGATAGGACGGACACATAAAGCTCACCGACTTTGGCCT  
834 I H R D I K P D N I L I D R D G H I K L T D F G L  
Intron 5

3676 GTGCACGGGATTCCGATGGACGCACAACCTCGAAGTACTACCAGGAGAACGGCAATCACTCGCGCCAGGACTCGAT  
859 C T G F R W T H N S K Y Y Q E N G N H S R Q D S M  
Intron 6

3751 GGAGCCCTGGGAGGAATACTCCGAGAACGGACCGAAGCCACCGTCTGCTGGAGAGGGACGGATGCCCGATCACCA  
884 E P W E E Y S E N G P K P T V L E R R R M R D H Q  
A

3826 AAGAGTCCTGGCCCACTCGCTGGTGGGCACCCGAAGTACATAGCTCCCGAGGTGCTGGAGAGGAGTGGGTACAC  
909 R V L A H S L V G T P N Y I A P E V L E R S G Y  
C T

3901 GCAGCTGTGCGACTACTGGAGCGTGGCGGTATCCTTTACGAGATGCTGGTGGGTGAGCCGCCCTTTCTGCCCA  
934 G L C D Y W S V G V I L Y E M L V G Q P P F L A A  
Intron 7 C

3976 CAGTCCGCTGGAACGCAACAAAAGGTCACTCAACTGGGAGAAAAGCTGCATATTCGCGCCAGGCGGAGTTAT  
959 S P L E T Q Q K V I N W E K T L H I P P Q A E L D

4051 CCGCGAGGCTACGGACTTGATAAGGAGGCTCTGTGCGTCCGCTGACAAGCGGCTGGCAAGAGCGTGGACGAGGT  
984 R E A T D L I R R L C A S A D K R L G K S V D E V

4126 CAAGAGCCACGACTTCTTCAAGGGCATCGACTTTGCGGACATGCGGAAGCAGAAAGCGCCCTACATACCGGAAAT  
1059 K S H D F F K G I D F A D M R K Q K A P Y I F E I

FIG.5C

BAD ORIGINAL

SUBSTITUTE SHEET (RULE 26)



13/43

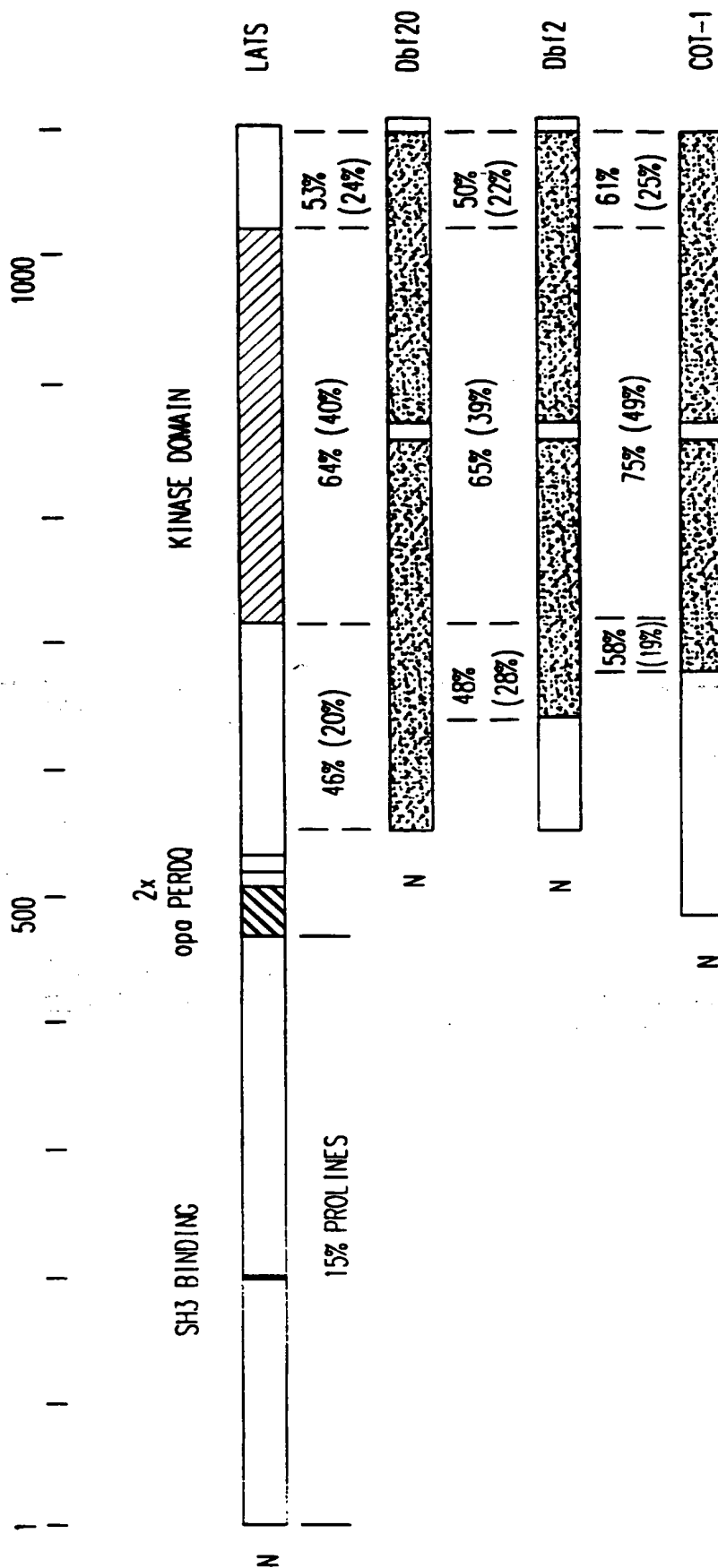


FIG.6A

LATS DROSOPHILA 546 SNNNNSEIKPPSPNNNNIQTSSNLAATIPP~~IP~~FAKYN-NSSNTQANSSGCSNCTGTIASSSTCKKI---KHA~~SP~~IPERK~~IK~~SK~~EE~~ERKEFRIRQTS  
 PK1L7 TOBACCO | MD~~S~~ARGWFCKLSSITKID-----E~~HA~~SGREDDKPVSAEASNTIKQV  
 PK SPINACH | MEK  
 DBF20 YEAST | MF~~S~~RSDREVD~~DL~~AG~~N~~SHLGFYDLNIR-K~~FI~~SP~~Q~~Q~~Q~~PPARKSENGRLTP~~Q~~PRSYKPCDSDDDDTF~~Q~~RI~~SL~~N~~SP~~KKL~~PK~~-DF~~ER~~ASQ~~S~~ITQ~~RV~~NNVC  
 DBF2 YEAST 82 ERATSNKIQ~~RV~~SVVC

LATS DROSOPHILA 644 POAF-KITFMECHIE~~N~~VIKSYRO-RTYRKNO-LE~~EM~~KVQLPQIQIEM---R~~Q~~LN~~CK~~ESNTIRLKR~~AK~~MDKSM~~FK~~PI~~Q~~GVGAFGEHMLVS-KIDTS  
 COT-1 NEUROSPORA 191 FROSE-M~~Q~~KLGETN-DARRRESI---W~~ST~~AGRKEGQY~~RF~~LR~~IK~~OKPENYQTI~~IK~~IGKGAFC~~EL~~VLQ-K-KAD  
 PK1L7 TOBACCO 43 AN~~Q~~ KQYIE~~K~~HYREOMKN-LOERRRRIIL-LE~~KK~~LADADVSEEDQNNL---L~~Q~~TE~~KK~~ETEM~~RL~~QR-K~~KG~~ADDE~~ELL~~TMIGCAFGEPTCMIGFSVITG  
 PK COMMON  
 ICE PLANT 1 RKLHDADVSEEDQNNL---L~~Q~~TE~~KK~~ETEM~~RL~~QR-K~~KG~~ADDE~~ELL~~TMIGCAFGEHVR-VC-REKJT  
 PK SPINACH 4 VK~~Q~~A-K~~Q~~ETIENHYR~~S~~OMKN-IQERKERWV-LE~~Q~~LASSDVPEEDQNSL---I~~Q~~LE~~FK~~ETEF~~RL~~K~~NR~~ICVND~~ELL~~TIIGCAFGEHVD-C-REKKS  
 DBF20 YEAST 100 QL~~Y~~LDYYCDM-FDYVI-SRRQ-RTIKOVL~~RY~~LEQ~~RS~~KNVSNKVLNEE---W~~AL~~YLQ~~RE~~HEVL~~PK~~RL~~PK~~PKH~~KD~~Q~~Q~~ITQVGGGYGO-WLAK-KQLS  
 DBF2 YEAST 97 KMYILEYYCDM-FDYVI-SRRQ-RTI--KQV-LE~~Y~~LQOQSQ~~PN~~SDQIKLNEEWSSYLQ~~RE~~HOVL~~PK~~RL~~Q~~PKNRDEMI~~TOV~~GGGYGO-WLAK-KQIT-

← KINASE DOMAIN

LATS DROSOPHILA 737 NII LYAMK~~TI~~IRKADVLKRNOVAHVKAERDILAE~~ND~~NNVVKLYYSFOQ~~Q~~NLY~~MD~~YIPGGDLMSLLIKLGTFFELARFYIAE~~VT~~CAVDSVH  
 COT-1 NEUROSPORA 259 CK VYAMKSLIKTEMFKKDQLAHVRAERDILAE~~SD~~SPHVVVKLYTTFQDANFLYMLMEFLPCCGLMMLIKYETESEDI~~TRF~~YIAEIVL~~AD~~NVH  
 PK1L7 TOBACCO 137 QNCREKITGOVYAMKLIKSEMRRGOVEHVKAERNLLAE~~ND~~SDCINKLYYSFOQ~~Q~~Q~~Q~~YLY~~ME~~YIPGGDMITLLMRKQITL~~TE~~DEARFYV~~AE~~ITVL~~AE~~SH  
 PK COMMON  
 ICE PLANT 63 GH VYAMKLIKSEMRRGOVEHVKAERNLLAE~~ND~~SNCINKLYQSFQ~~Q~~EEYLY~~IN~~EMYPGGDMITLLMRKQITL~~TE~~DEARFYV~~AE~~ITVL~~AE~~SH  
 PK SPINACH 96 EN VYAMKLIKSEMRRGOVEHVKAERNLLAE~~ND~~SHCINKLYFSFOQ~~AE~~YLY~~IN~~EMYPGGDMITLLMRKQITL~~TE~~KV~~AG~~YIAQSVL~~AE~~SH  
 DBF20 YEAST 193 DE ICALKILNKKLLFKLNEIN~~VL~~TERDILITTRSD~~ML~~NKLLMAFQ~~DP~~ESYLY~~AN~~EFVPGGDFRTLLIN~~TR~~ILKSGHARFYISE~~AF~~CAVNA~~LI~~H  
 DBF2 YEAST 190 KE VCAKILNKKLLFKLNEIKH~~ML~~TERDILITTRSE~~ML~~NKLLMAF~~Q~~LOSLYLY~~AN~~EFVPGGDFRTLLIN~~TR~~CLKSGHARFYISE~~AF~~CAVNA~~LI~~H

FIG.6B

15/43

LATS DROSOPHILA 830 KMGTHIRDIKPDNLLIDRGCHIKLIDFGLCTGIRW--THNSKYDENG--NHSRODSMEPWEYSE--PKPIT--VLERRM--RDHORM--ASLNGTIPNYIAP  
 COT-1 NEUROSPORA 352 KLGFTHIRDIKPDNLLIDRGCHVKLIDFGLSTGTHK--LJDNWYITQLL--QGYSNKPDRNRSVAIDQINLL--VSNPAWINDWRRSRRLMAYSTINGTPDYIAP  
 PKT17 TOBACCO 238 KINYTHIRDIKPKNLLIDRGCHKLSDGFLCHPLDCSTILEEKF--SWG--DNANGGSRSDSPP--APKRTQOE--QLEHWQ--KNR--RMAYSTINGTPDYIAP  
 PK COMMON  
 ICE PLANT 156 KINYTHIRDIKPDNLLIDKFGHLRLSDFGLCMPDCSTILEEKF--VNN--GNGGSPSLEESTK--PRR--TQOE--QLEHWQ--KNR--RMAYSTINGTPDYIAP  
 PK SPINACH 189 KINYTHIRDIKPDNLLIDKNGHMLSDFGLCMPDC--ATLSTIKENE SMOVDYSKNSMDIDASLPDAGCHSWPSAREQLQWQNRRLAFSTINGTPDYIAP  
 DBF 20 YEAST 285 ELGYTHRDLPENFLIDATGHIKLLIDFGLAGCT--VSNRIESMKI--RLEEVKNLQ--FPATERSIEDR--SKYHNM--RKTEINYNANSGSPDMAL  
 DBF 2 YEAST 282 DLGYTHRDLPENFLIDAGHIKLLIDFGLAGCT--ISNRIESMKI--RLEKIKOLE--FPATERSIEDR--RKMYNQL--REKEINYNANSGSPDMAL

## KINASE DOMAIN

LATS DROSOPHILA 925 LVIERSGYTQLCDYWSGCVILYEMLVGPPELANSPLETDQVINEKTLHIP---PQAFSPEADLT--PRLCASADKRLG--KSDVEKSHDFKGIIDE  
 COT-1 NEUROSPORA 450 EFTGICYSTCDWWSLGTILMECLVGMPPECAEDSHDIYRKIVNPHSLYFP---DDITLGVDAWNL--PRLICNTENRLCGGAHEIKSHAFFRGVEF  
 PKT17 TOBACCO 330 EVIILKKGYGMECDWWSLCAIMYEMLVGPPEYSDDPMSICRKIVNKNHKKFP---EEAKLSPEAKDILISLLCNVTE--RLGSNCADEIKKASWFGIDW  
 PK COMMON  
 ICE PLANT 248 [VLILKKGYGMECDWWSLCAIMYEMLVGPPEYSDDPMSICRKIVNPHTHKKFP---EEAKLSPEAKDILISLLCNVTO--RLGSNCAHEIKLFWNGIDW  
 PK SPINACH 289 EVILKKGYGMECDWWSLCAIMYEMLVGPPEYSDDPITICRKIVNPHYIKFP---DDAKLFEADLIDRLLC--DVEHRLGTGAEOIKVHAWKDEW  
 DBF 20 YEAST 377 [VIEGKKYDFIVDWSLGCMEFSLVGYTFPSSSSINEIYENRYNKKTLRPRRTEDRAAFSDRIWDL--TRLIADPINRV--RSTEQRMKMSYFAELNF  
 DBF 2 YEAST 374 [VIEGKKYDFIVDWSLGCMEFSLVGYTFPSSSSINEIYDNLRYNKKTLRPRROSDGGAFAFSDRIWDL--TRLIADPINRL--RSTEHMKRMSYFAELNF

LATS DROSOPHILA 1020 ADMRKOKAPYIPEIKIHTIDISNFDPMVPEKLRNDISIMSSGDVDOQ-----ND-RTHGFEEFTFRFF---DDKQTFDMIDDOAPVYV.  
 COT-1 NEUROSPORA 546 DSIERRIRAPFEPRLISAIDITYF--PID--EIDQIDNAIILKAQQAARGAAPAQGEESPELSPTIGYTFRRF-----DNIFR.  
 PKT17 TOBACCO 426 DRIYQMEAAFPEVNDDELIDNFEKTE--ESESHSGSGSRSGCPWRKML-----SSKDNFVGYTYKNF--KVNDQVQFDMVLMELKKTNIK + 20 A.A.  
 PK COMMON  
 ICE PLANT 344 [LRIYQMEAAFPEVNDDELIDNFEKTE--EADNSQSISKAGCPWRKML-----SSKDLNFGVITYKNFEIVNDQVPGIAELKKKQIK + 55 A.A.  
 PK SPINACH 385 DRIYEDNAYKPOVNGELIDNFMKFD--EANPPTPSRSGCPWRKML-----ISKDLSFVGYTYKNF--DAVKGLKHSFDRKGSTS + 39 A.A.  
 DBF 20 YEAST 475 [LIRISSPPEIPQDDEITDAGYFDDFTNEEDMAKYADVFQRONKLSAMV-----DDSAVDSKLVGFTFR--DDKQSSGILYNGSEIHSDFSTFYD.  
 DBF 2 YEAST 472 SITIRSMIIPPEIPQDSEITDAGYFDDFTSEIDMAKYADVFQRONKLTAMV-----DDSAVSSKLVGFTFR--DDKQSSGILFNGLEHSDFSTFYD.

FIG. 6C

16/43

10	20	30	40	50	60	70	80
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *
GIGGACAT	CAATAC	AAACAAGC	TGGMAGGT	CTAAGAGIC	TCIAGTICCT	CAGAGACAG	GCCCCATCCT
V Q H S	I N R K	Q S W K	G S K E	S L V P	Q R H G	P S L	160
100	110	120	130	140	150	160	
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	
AGGAGAAAT	GIGGTTAIC	GTTCIGAVG	CCCCAACTCA	CAGGGGGAIG	TAGGAAGACC	TCTGICTGGA	TCCGGCATTG
G I N V	V Y R S	E S P N	S Q A D	V G R P	L S G S	G I	
170	180	190	200	210	220	230	240
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	
CAGCATTCG	TCAGCTCAC	CCAAGCAATG	GACAGAGAGT	GAACCCCCCA	CCACCACCTC	AAGTTAGGAG	TGTTACTCCT
A A I A	Q A H P	S N G Q	R V N P	P P P P	Q V R S	V T P	
250	260	270	280	290	300	310	320
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	
CCACCACTTC	CGAGAGGCCA	GACCCCACT	CCCCGAGGCA	CCACTCCCCC	TCCCCCTCA	TGGGAACCAA	GCTCTCAGAC
P P P P	P R G Q	T P P P	P R G T	T P P P	P P S W	E P S S	Q T
330	340	350	360	370	380	390	400
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	
AAAGGGLAC	ICTCTGAA	IGGAGIACGT	AATCICCGA	ATCTCCCCIG	TTCCACCTGG	GGCGTGGCAG	GAGGGGTACC
K R Y S	G N M E	Y V I S	R I S P	V P P G	A W Q E	G Y	
410	420	430	440	450	460	470	480
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	
CICCAACAC	ICTTACCACI	ICTCCCAIGA	ATCCCCCTAG	CCAGGCTCAG	AGGGCCATTA	GTTCIGTCC	AGTTGGTAGA
P P P P	I I I S	P M N P	P S Q A	Q R A I	S S V P	V G R	
490	500	510	520	530	540	550	560
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	
CAACCAATCA	TCATGCAGAG	TACTAGCAAA	TTTACTTTA	CACCAGGGCG	ACCTGGAGTT	CAGATGGTG	GTGGTCAGTC
Q P I I	H Q S I	S K F N	I T P	G R P G	V Q N	G G G Q S	

FIG. 7A

BAD ORIGINAL



17/43

570	*	580	*	590	*	600	*	610	*	620	*	630	*	640	*
TGATTTTATC GGGCACCAMA ATGICCCCCAC TGGTTCTGTG ACTCGGGCAGC CACCACCCTCC ATATCCTCTG ACCCCAGCTA															
D F I V H Q N V P T G S V T R Q P P P P Y P L T P A															
650	*	660	*	670	*	680	*	690	*	700	*	710	*	720	*
AAGGACAAAG CCCCCTCTGCT TTACAAACAG GGGCTTCTGC TGCTCCACCA TCATTGCGCA ATGGAACGT TCCTCAGTCG															
N G Q S P S A L Q T G A S A A P P S F A N G N V P Q S															
730	*	740	*	750	*	760	*	770	*	780	*	790	*	800	*
ATGAIGGTC CCAACAGGAA CAGTCATAAC ATGGAGCTTT ATAATATTAA TGTCCTTGA CTGCAACAG CCTGGCCCCA															
M M V P N R N S H N M E L Y N I N V P G L Q T A W P Q															
810	*	820	*	830	*	840	*	850	*	860	*	870	*	880	*
GTGGTCTCT GTGCTCTGGC AGTCATCCCC AAGCGGTGGG CATGAAATTC CTACATGGCA ACCTAACATA CCAGTGAGGT															
S S S A P A Q S S P S G G H E I P T W Q P N I P V R															
890	*	900	*	910	*	920	*	930	*	940	*	950	*	960	*
CAATCTCTT TATTAACCA TTAGGAGTGA GAGCMGTCA CTCTGTCTAT TCICAGCCTT CTGCCACTAC AGTCACCTGCC															
S N S I H N P L G S R A S H S A N S Q P S A T T V T A															
970	*	980	*	990	*	1000	*	1010	*	1020	*	1030	*	1040	*
ATCACACCG CTTCTATICA ACAGCCCGTG AAAAGCATGC GCGTCTGAA ACCAGAGCTG CAGACTGCTY TAGCCCCAAC															
I T P A P I Q Q P V K S M R V L K P E L Q T A L A P T															
1050	*	1060	*	1070	*	1080	*	1090	*	1100	*	1110	*	1120	*
CCAATCTCT TCAATGCCAC AGCCAGTICA GACTGTICAG CCTACCCCTT TTTCIGAGGG TACAGCTTCA AGTGTGCTCTG															
H P S H H P Q P V Q I V Q P T P F S E G T A S S V P															

FIG. 7B

BAD ORIGINAL

18/43

1130	1140	1150	1160	1170	1180	1190	1200
*	*	*	*	*	*	*	*
TCATCCCACC	IGITGCTGAA	GCITCCAAGCT	ATCAAGGTCC	ACCACCGCCT	TATCCAAAAC	ATCTGCTACA	CCAAAACCGA
V I P P	V A E	A P S	Y Q G P	P P P	Y P K	H L L H	Q N P
1210	1220	1230	1240	1250	1260	1270	1280
*	*	*	*	*	*	*	*
TCTGICCCCTC	CATATGAGTC	AGTAAGTAAG	CCCTGCAAG	ATGAACAGCC	TAGCTTACCC	AAGGAAGATG	ATAGTGAGAA
S V P	P Y E S	V S K	P C K	D E Q P	S L P	K E D	D S E K
1290	1300	1310	1320	1330	1340	1350	1360
*	*	*	*	*	*	*	*
GAGTCCGGAC	AGTGGTGACT	CTGGGGATAA	AGAAAAGAAA	CAGATTACAA	CTTACACCTAT	CACGTGTCGG	AAAAACAAAG
S A D	S G D	S G D K	E K K	Q I T	T S P I	T V R	K N K
1370	1380	1390	1400	1410	1420	1430	1440
*	*	*	*	*	*	*	*
AAGAIGMGA	ACGAMGAGAG	ICTCGGATTC	AGAGTTACTC	CCCACAGGCC	TTTAAAGTTCT	TCATGGAGCA	GCACGTAGAG
K D E	F R R E	S R I	Q S Y S	P Q A	F K F	F M E	Q H V E
1450	1460	1470	1480	1490	1500	1510	1520
*	*	*	*	*	*	*	*
AAGGICCCIGA	AGTICATCA	GCAGCGTCTG	CATCGGAAGA	AGCAGCTAGA	AAATGAAATG	ATGCGGGTTG	GATTATCTCA
N V I	K S H Q	Q R L	H R K	K Q L E	N E M	M R V	G L S Q
1530	1540	1550	1560	1570	1580	1590	1600
*	*	*	*	*	*	*	*
AGAIGCCGAG	GATCAAAIGA	GAAGAIGCT	TTGCCAGAAA	GAGTCTAACT	ATATTCGCTCT	TAAAGGGCT	AAAATGGACA
D A Q	D Q M	R K M L	C Q K	E S N	Y I R L	K R A	K M D
1610	1620	1630	1640	1650	1660	1670	1680
*	*	*	*	*	*	*	*
AGTCTAIGIT	IGIAAGAAIA	ANGACAT IAG	GAATAGGAGC	GTTTGGIGAA	GICIGIC IAG	CAAGAAAAGT	CGATACTAAA
K S H I	V K I	K T I	G I G A	F G E	V C L	A R K	V D T K

FIG. 7C

BAD ORIGINAL

19/43

1690	*	1700	*	1710	*	1720	*	1730	*	1740	*	1750	*	1760	*
GCCTTGTG	CGACAAAGAC	TCCTCGAAAG	AAAGACGTTT	TGCTCCGAAA	TCAGGTTGGT	CACTGAAAG	CGGAGAGGGA								
A L Y A T K I L R K K D V L L R N Q V A H V K A E R D															
1770	*	1780	*	1790	*	1800	*	1810	*	1820	*	1830	*	1840	*
TAATCCAGCA	GAGCCGACA	AAGAGTGGT	GGTCCGCTG	TACTACTCT	TCCAGGACAA	GGACAACTTG	TACTTTGTGA								
I L A I A D N E W V V R L Y Y S F Q D K D N L Y F V															
1850	*	1860	*	1870	*	1880	*	1890	*	1900	*	1910	*	1920	*
TGGACTACAT	TCCTGGGGG	GATATGATGA	GCCTATTAAAT	TAGAATGGC	ATCTTTCCTG	AAAATCTGGC	ACGATTCTAC								
M D Y I P G G D M M S L L I R M G I F P E N L A R F Y															
1930	*	1940	*	1950	*	1960	*	1970	*	1980	*	1990	*	2000	*
ATAGCAGAAC	TAACCTGIGC	AGTTGAAAGT	GTTCAATAA	TGGGTTTTAT	TCATAGAGAT	ATTAAACCTG	ATAACATTTT								
I A L I I C A V E S V H K M G F I H R D I K P D N I L															
2010	*	2020	*	2030	*	2040	*	2050	*	2060	*	2070	*	2080	*
GATTGACCTT	GATGGCCAIA	IAAAATGAC	TGACITGGC	TGTGACCTG	GCTTCAGATG	GACACATGAC	TCCAAGTACT								
I D R D G H I K L T D F G L C T G F R W T H D S K Y															
2090	*	2100	*	2110	*	2120	*	2130	*	2140	*	2150	*	2160	*
ACCAGAGTGG	GGATCACCCA	CGGAAGATA	GCATGGATT	CAGTAACGAA	TGGGGAGATC	CTTCCAATTG	TCGGTGTGGG								
Y Q S G D H P R Q D S M D F S N E W G D P S N C R C G															
2170	*	2180	*	2190	*	2200	*	2210	*	2220	*	2230	*	2240	*
GACAGACTGA	AGCCACTGGA	GCGGAGAGCT	GCTCGCCAGC	ACCAGCGATG	TCTAGCCCAT	TCTCTGGTTG	GGACTCCCAA								
D R I K P I T R R A A R Q H Q R C L A H S L V G T P N															

FIG. 7D

BAD ORIGINAL

20/43

2250 \* 2260 \* 2270 \* 2280 \* 2290 \* 2300 \* 2310 \* 2320 \*  
 TTAATGCA CCGACAGGAC IACIGCGAAC AGGATATACA CAGCTGTGTG ACTGGGAG TGTGGIGTT ATTCTTTGTG  
 Y I A P I V I I R I G Y T Q L C D W S V G V I L C  
 2330 \* 2340 \* 2350 \* 2360 \* 2370 \* 2380 \* 2390 \* 2400 \*  
 AAAAGTGGI GCGACACCT CCTTCTTGG CACAAACCCC AITAGAACAA CAAATGAAGG TTAATCACTG GCAAACTTCT  
 E M L V G Q P P F L A Q T P L E T Q M K V I I W Q T S  
 2410 \* 2420 \* 2430 \* 2440 \* 2450 \* 2460 \* 2470 \* 2480 \*  
 CTACACATCC CCGCICAGC TAAGCIGAGT CCTGAAGCCT CTGACCTCAT TATCAAACTG TGTGAGGAC CAGAAGACCG  
 L H I P P Q A K L S P E A S D L I I K L C R G P E D R  
 2490 \* 2500 \* 2510 \* 2520 \* 2530 \* 2540 \* 2550 \* 2560 \*  
 CCICGGCAGG AAGGIGGICG AIGAGATAAA GGCATATCCA TTTTAAAG CCAATCGATT CTCTAGTAT CTGAGACAGC  
 L G K H G A D E I K A H P F F K T I D F S S D L R Q  
 2570 \* 2580 \* 2590 \* 2600 \* 2610 \* 2620 \* 2630 \* 2640 \*  
 AGICTGCTTC ATACATCCCI AAAATCAGC ATCCAACAGA TACATCCAAT TTCGACCCCTG TTGATCCIGA TAAATGTGG  
 Q S A S Y I P K I T H P I D T S N F D P V D P D K L W  
 2650 \* 2660 \* 2670 \* 2680 \* 2690 \* 2700 \* 2710 \* 2720 \*  
 AGCGATGCTA GCGAGGAGGA AAATACAGT GACACTCIGA GCGGATGGTA TAAAAATGGG AAGCACCCCG AGCAGCGTTT  
 S D G S I L L N I S D T L S G W Y K N G K H P E H A F  
 2730 \* 2740 \* 2750 \* 2760 \* 2770 \* 2780 \* 2790 \* 2800 \*  
 CTATGAGCTT ACTTTTTCGA GGTHTTTCGA TACCCATATA ATTAATCCAA GCCTATIGAG TATGAATACA  
 Y I I I I R P I I I D D N G Y P Y N Y P K P I E Y E Y

FIG. 7E

BAD ORIGINAL

21/43

```

2810 * 2820 * 2830 * 2840 * 2850 * 2860 * 2870 * 2880 *
TTCATTCACA GGGCICAGAA CAACAGICTG ATGAAGATGA TCAACACACA AGCTCCGATG GAAACAACCG AGATCTAGTG
I H S Q G S E Q Q S D E D D Q H T S S D G N N R D L V
2890 2900 2910 2920 2930 2940 2950 2960
* * * * * * * *
TATGTTTAAI AACTAGGAG ATCATGTGAA GAAITTGCAA GAGGCCCTGAA GTGCAGGGGT TTTTGAAGTT TTGAGARAAT
Y V *
2970 2980 2990 3000 3010 3020 3030 3040
* * * * * * * *
TATGCAAAAG IGACAGAGTT TGTGTGCTCT GTGTACAATA TTTTATTTTC CIAAGTTATG GGAAATTGTT TTAAAAATGTT
3050 3060 3070 3080 3090 3100 3110 3120
* * * * * * * *
AATTTATTC ACCCTTTTAA TTCAAGTAAT TAGAAAAAAT TGTATAAGG AAAGTAAATT ATGAACTGAG TATTATAGTC
3130 3140 3150 3160 3170 3180 3190 3200
* * * * * * * *
AATCTCTGGT ACTTAAAGTA CTTAAAAGA GAAGCCCTGGT AICTTTTGTG TATATAATAA ATAATTTTAA AATCCCAAAA
3210
*
AAAAAAAA AA

```

FIG. 7F

BAD ORIGINAL

22/43

10	20	30	40	50	60	70	80
* ATGAGAGCCA	* CCCCGAGGTT	* IGGACCTTAT	* CAAAAAGCTC	* TCAGGGAAT	* CCGATATTCC	* CTCCTGCCCT	* TTGCCAACGA
90	100	110	120	130	140	150	160
* M R A	* I P K F	* G P Y	* Q K A	* L R E I	* R Y S	* L L P	* P A N E
170	180	190	200	210	220	230	240
* S G I	* S A A	* A E V N	* R Q M	* L Q E	* L V N A	* A C D	* Q E M
250	260	270	280	290	300	310	320
* A G R A	* I I Q	* I G S	* R S I E	* A A L	* E Y I	* S K M	* G Y L D
330	340	350	360	370	380	390	400
* CCCAGGATG	* AGAGATIG	* GGGAGTATC	* AAGCAGACCT	* CCCCAGGAAA	* GGGCTGGCG	* TCCACCCCGG	* TGACTCGGCG
410	420	430	440	450	460	470	480
* P S F	* I G I	* G F A L	* P S Y	* H Q L	* G G A N	* Y E G	* P A A
490	500	510	520	530	540	550	560
* TGGAGGAGAT	* GCGCGGCCAA	* IATTTAGACT	* TTCTCTTCCC	* TGGAGCCGGA	* GCCGGCACCC	* ACGGTGCCTA	* GGCTCACCAG
570	580	590	600	610	620	630	640
* L E E	* H P R	* Q Y L	* D F L	* F P G	* A G A	* G T H	* G A Q
650	660	670	680	690	700	710	720
* CATCTCTCCA	* MGGGTACAG	* CACAGGCAGIA	* GAGCCAAGTG	* CGCACTTTCC	* GGGCACACAC	* TATGGTCGTG	* GTCATCTACT
730	740	750	760	770	780	790	800
* H P P	* R G Y S	* I A V	* E P S	* A H F	* P G T	* H Y G	* R G H L L

FIG. 8A

BAD ORIGINAL

23/43

570	*	580	*	590	*	600	*	610	*	620	*	630	*	640	*
ATCGGAGCAG	TCIGGGIATG	GGGTGCAGCG	CAGTTCCTCC	TTCCAGAACA	AGACGCCACC	AGATGCCCTAT	TCCAGCATGG								
S E Q S G Y G V Q R	S S S F Q N K T P P	D A Y S S M													
650	*	660	*	670	*	680	*	690	*	700	*	710	*	720	*
CCAAGGCCCA	GGGIGGCCCT	CCGGCCAGCC	TCACCTTICC	TGCCCATGCT	GGGCTGTACA	CTGGCTCGCA	CCACAAGCCG								
A K A Q G G P P A S	L T F P A H A G L Y	T A S H K P													
730	*	740	*	750	*	760	*	770	*	780	*	790	*	800	*
GGGGTACCC	CACCIGGGGC	CCACCCATTA	CATGTGTGG	GCACCGGGG	TCCCACGTTT	ACTGGCGAAA	GCTCTGCACA								
A A I P P G A H P L	H V L G T R G P T F	T G E S S A Q													
810	*	820	*	830	*	840	*	850	*	860	*	870	*	880	*
GGCTIGCTIG	GCACCGICCA	GGMACAGCT	CAATGCIGAC	TTGTACGAGC	TGGGCTCCAC	GGTGCCCCIGG	TCITGCAGCTC								
A V L A P S R N S L	N A D L Y E L G S T	V P W S A A													
890	*	900	*	910	*	920	*	930	*	940	*	950	*	960	*
CACIGGCACG	CCGGACICG	CIGCAGAAGC	AGGGTCTAGA	AGCCTCGCGG	CCGCAATGTGG	CTTTCGGGC	TGGCCCCCAGC								
P L A R R D S I Q K	Q G L E A S R P H V	A F R A G P S													
970	*	980	*	990	*	1000	*	1010	*	1020	*	1030	*	1040	*
AGGACCAACT	CCCTTCACMA	CCCACAACCT	GAGCCCTCAC	TGCCCGCCCC	CAACACGGTC	ACCGCCGTGA	CGGCCGCACA								
R T N S F N N P Q P	E P S L P A P N T V	T A V T A A H													
1050	*	1060	*	1070	*	1080	*	1090	*	1100	*	1110	*	1120	*
CACTCTTACG	CCCTIGCAGA	GCGTGGGTGT	GCTGGCGCCC	GAGCCCCAGA	CAGCCGTGGG	GCCCTCGCAC	CCCGCCTGGG								
I I H P V K S V R V	L R P E P Q T A V G	P S H P A W													

FIG. 8B

BAD ORIGINAL

24/43

1130	1140	1150	1160	1170	1180	1190	1200
* GGCTGGGTC	* CAGACGACCT	* GCCACIGAGA	* GCCIGGAGAC	* GAAGAGGGC	* AGCGCAGGCC	* CACACCCGCT	* GGATGTGGAC
V A A P I A P A T E S L E T K E G S A G P H P L D V D	1220	1230	1240	1250	1260	1270	1280
* TAAGGGGCT	* CCGAGCGCAG	* GIGCCACCG	* CCTCCGATC	* CAAAGCACTT	* GCTGCTGCC	* AGTAAGTCTG	* AGCAGTACAG
Y G G S I R R C P P P P Y P K H L L L P S K S E Q Y S	1300	1310	1320	1330	1340	1350	1360
* CGGGGCTG	* GACAGCCCTG	* GCACCAGTGT	* GCAGCAGAGT	* CTGGAGGGG	* GCACGTATCT	* AGACGGGAGT	* GACAAGAGCC
V D L D S L C T S V Q Q S L R G G T D L D G S D K S	1380	1390	1400	1410	1420	1430	1440
* ACAAAGGTG	* GAAAGGAGAC	* AAGCTGGCA	* GAGACAAAA	* GCAGATTAC	* ACCTCCCGG	* TGCCTGTCCG	* CAAGAATAGC
H K G A F G D K A G R D K K Q I Q T S P V P V R K N S	1460	1470	1480	1490	1500	1510	1520
* AGAGATGAG	* AGAAAGAGA	* GICTCGCATC	* AAGAGTAC	* CCCCTTATGC	* CTTCAAATTC	* TTCAIGGAGC	* AACACGTGGA
R D L I K R E S R I K S Y S P Y A F K F F M E Q H V E	1540	1550	1560	1570	1580	1590	1600
* GATGTGATC	* AAACCTTACC	* AGCAGAAGGT	* CAGCCGGAGG	* CTACAGCTGG	* AGCAGGAAT	* GGCCAAGCT	* GGGCTCTGTG
N V I K I Y Q Q K V S R R L Q L E Q E M A K A G L C	1620	1630	1640	1650	1660	1670	1680
* AGGCCGATTA	* GAGGAGATG	* AGGAAGATCC	* TCTACCAGAA	* GGAGICTAAC	* TACAACCGGC	* TGAAGAGGGC	* CAAGAATGGAC
F A I O I Q M R K I I Y Q K E S N Y N R L K R A K M D							

FIG. 8C

BAD ORIGINAL



25/43

1690	*	1700	*	1710	*	1720	*	1730	*	1740	*	1750	*	1760	*
AMGICCAIGI	IGIGAAAT	CAAGACICIA	GGCATCGGTG	CCTTIGGGGA	AGIGIGCCIC	GCTTIGIAAGC	TGGACACTCA								
K S H	F V K I	K T L	G I G	A F G E	V C L	A C K	L D T H								
1770	*	1780	*	1790	*	1800	*	1810	*	1820	*	1830	*	1840	*
CGCTCIGTAC	GUCAIGAAGA	CICTCAGGAA	GAAGGAITC	CTGAACCGGA	ATCAAGIGGC	CCAIGTCAAG	GCTGAGAGGG								
A L Y	A M K	I L R K	K D V	L N R	N Q V A	H V K	A E R								
1850	*	1860	*	1870	*	1880	*	1890	*	1900	*	1910	*	1920	*
ACAICCTGGC	IGAAGCAGAC	ATGAGIGGG	TGGTCAACT	CTACTACTCC	TTCCAGGACA	AGGACAGCCT	GTACTTIGTG								
D I L	A F A D	N E W	V V K L	Y Y S	F Q D	K D S L	Y F V								
1930	*	1940	*	1950	*	1960	*	1970	*	1980	*	1990	*	2000	*
ATGGACTAAC	IACCAGGGG	GGATAIGATG	AGCCTGCTGA	TCAGGATGGA	GGTCTTCCCT	GAGCACCTGG	CCGGCTICIA								
M D Y	I P G G	D M M	S L L	I R M E	V F P	E H L	A R F Y								
2010	*	2020	*	2030	*	2040	*	2050	*	2060	*	2070	*	2080	*
CATTGCAAG	IGACCCGIG	CCATIGAAG	IGTCCACAG	ATGGGCTTIA	ICCACCAGGA	CATCAAGCCT	GACAACATAC								
I A I	I I I	A I E S	V H K	M G F	I H R D	I K P	D N I								
2090	*	2100	*	2110	*	2120	*	2130	*	2140	*	2150	*	2160	*
TCAICCACT	GGAGGICAT	ATTAAGCIGA	CAGATTIIGG	CCICIGCACT	GGATTACAGT	GGACTCACAA	TTCACAGTAC								
L I D	I D G H	I K L	T D F G	L C T	G F R	W T H	N S K Y								
2170	*	2180	*	2190	*	2200	*	2210	*	2220	*	2230	*	2240	*
IACCAGAAAG	GGACACACAI	GAGACAGGAC	AGCATGGAGC	CCGGTIGACCT	CTGGGACGAT	GTTTCCAACT	GTCGCTIGGG								
Y Q E	G H H N	R Q D	S M I	P G D L	W D D	V S N	C R C G								

FIG. 8D

BAD ORIGINAL

26/43

2250	*	2260	*	2270	*	2280	*	2290	*	2300	*	2310	*	2320	*
AGACAGGTTA	AGACAGGTTA	AGCAGAGGGC	GCAGAAGCAG	CACCAGAGGT	GCCTGGCACA	TTCTCTIGTC	GGGACACCAA								
D R L K I L F Q R A Q K Q H Q R C L A H S L V G T P															
2330	*	2340	*	2350	*	2360	*	2370	*	2380	*	2390	*	2400	*
ATTACATCGC	TCUGGAGGIG	CITCICCGCA	AAGGTACAC	GCAGCICIGT	GACTGGTGGG	GCGTCGGTGT	GATTCICITTT								
N Y I A P E V L L R K G Y T Q L C D W W S V G V I L F															
2410	*	2420	*	2430	*	2440	*	2450	*	2460	*	2470	*	2480	*
GAGATGCTGG	TTGGGCAGCC	GCCTTTCTIG	GCCCCCAGCC	CCACAGAGAC	GCAGCTGAAG	GTGATCAACT	GGGAGAGGCAC								
E M L V G Q P P F L A P T P T E T Q L K V I N W E S T															
2490	*	2500	*	2510	*	2520	*	2530	*	2540	*	2550	*	2560	*
GCTGGATATC	CCACAGCAGG	IGAGGCTCAG	CGCTGAGGCC	CGAGACCICA	TCACGAAGCT	GTGCTGGCGG	GCTGACTGCC								
L H I P I Q V R L S A E A R D L I T K L C C A A D C															
2570	*	2580	*	2590	*	2600	*	2610	*	2620	*	2630	*	2640	*
GCTTGGGAG	CGATGGGGCA	GATGACCICA	AGGCACACCC	GTTCCTCAAC	ACCATCGACT	TTTCCCGTGA	CATCCGAAAG								
R L G R D G A D D L K A H P F F N T I D F S R D I R K															
2650	*	2660	*	2670	*	2680	*	2690	*	2700	*	2710	*	2720	*
CAGGCTGGAC	CCACGCTGCC	CACCATCAGC	CACCCCAIGG	ACACCCTCAA	TTTIGACCCG	GTGGAIGAAG	AAAGCCCCCIG								
Q A A P Y V P I I S H P M D T S N F D P V D E E S P W															
2730	*	2740	*	2750	*	2760	*	2770	*	2780	*	2790	*	2800	*
GCACGAGCTC	ATGCGAGAGA	GCGCCAGGC	CTGGGACAGC	CTGGGCTCCC	CCAGCAGCAA	GCATCCAGAG	CACGCCCTCT								
H I A S G T S A K A W D T L A S P S S K H P E H A F															

FIG. 8E

BAD ORIGINAL

27/43

2810	*	2820	*	2830	*	2840	*	2850	*	2860	*	2870	*	2880	*
ATGAGTTCAC CTTCCGAGG TCTTCGATG ACAACGGCTA TCCCTCCGG TGCCCGAAGC CCTCAGAGCC CGCAGAGAGT															
Y E F I F R R F F D D N G Y P F R C P K P S E P A E S															
2890	*	2900	*	2910	*	2920	*	2930	*	2940	*	2950	*	2960	*
GCAGACCCAG GGAIGCGGA CTGGAAGGT GCGGCCGAGG GCTGCCAGCC GGTGTACGTG TAAGCCCTCAG TTAACCACAA															
A D P G D A D L E G A A E G C Q P V Y V *															
2970	*	2980	*	2990	*	3000	*	3010	*	3020	*	3030	*	3040	*
CTCGAGGAA CCCAAAATGA GATTCTTTT CAGAAGACAA ACTCAAGCTT AGGAATCCTT CATTITTAGT TCTGGTAAAT															
3050	*	3060	*	3070	*	3080	*	3090	*	3100	*	3110	*	3120	*
GGGCAACAGG AAGAGTCAAC ATGATTTCAA ATTAGCCCTC TGAGGACCTT CACTGCATTA AAACAGTATT TTTTAAAAAA															
3130	*	3140	*	3150	*										
TTAGTACAGT ATGGAAAGAG CACTTATTTT GGGGG															

FIG. 8F

BAD ORIGINAL

28/43

```

10      *      20      30      40      50      60      70      80      *
*      *      *      *      *      *      *      *
ACCTTGGG  IGC TGGGACG GACTCTGGCC GCCTCAGCGT CCGCCCTCAG GCCCGTGGCC GCTGTCAGG AGCTCTGCTC
90      *      100     110     120     130     140     150     160     *
*      *      *      *      *      *      *      *
TCCCCCAGC AGTAAATAT TATATTGTA AAGAAATTTA ACAGTCTCTGG GGAATTCCTT GAAGGATCAT TTTCACATTTT
170     *      180     190     200     210     220     230     240     *
*      *      *      *      *      *      *      *
GCTCAGAGGA AGC TCTGGA TCTATCAAT AAAGAAGTCC TTCGTGTGGG CTACATATAT AGATGTTTTT ATGAAGAGGA
M K R

250     *      260     270     280     290     300     310     320     *
*      *      *      *      *      *      *      *
GIGAAAGCC AGAAGGATAT AGACAAATGA GGCCTAAGAC CTTTCCTGCC AGTAACTATA CTGTCAGTAG CCGGCAAAATG
S E K P I G Y R Q M R P K T F P A S N Y T V S S R Q M
330     *      340     350     360     370     380     390     400     *
*      *      *      *      *      *      *      *
TTACAGGAA TCGGCAAT CTTAGGAAT TTAICTAAT CATCTGATC TGCTAAGGCT GAGCATAACA TGAGTAAAT
L Q L I R L S I R N L S K P S D A A K A E H N M S K M
410     *      420     430     440     450     460     470     480     *
*      *      *      *      *      *      *      *
GTCAGGAA GATCCCTGAC AGICAGAA TCCACCCAA TTTGGGACG ATCATAAAGC CTTCGAGGAA ATTCGAACT
S I L D P R Q V R N P P K F G T H H K A L Q E I R N
490     *      500     510     520     530     540     550     560     *
*      *      *      *      *      *      *      *
CTCAGCTCC ATTGCAAT GAAACAAT CTTCTGGAG TACTTCAGAA GTTAATCCAC AAATGCTTCA AGACTTGCAA
S L I P F A N E T N S S R S T S E V N P Q M L Q D L Q

```

FIG. 9A

BAD ORIGINAL

29/43

570	*	580	*	590	*	600	*	610	*	620	*	630	*	640	*
GCIGCTGAT	IGAIGAGGA	TAIGGTTATA	CAAGCTCTTC	AGAAACTAA	CAACAGAAGT	ATAGAAGCAG	CAATTGAATT								
A A G I D E D	M V I Q A L	Q K T N	N R S I E A	A I E F											
610	*	620	*	630	*	640	*	650	*	660	*	670	*	680	*
CAIAGTAA	ATGAGTTACC	AAGATCCTCG	ACGAGAGCAG	ATGGCTGCAG	CAGCTGCCAG	ACCTATTAA	GCCAGCATGA								
I S K M S Y	Q D P R	R E Q M A A	A A A R	P I N A S M											
730	*	740	*	750	*	760	*	770	*	780	*	790	*	800	*
ANCCAGGGA	IGIGCAGCAA	TCAGTTAACC	GCAACACAGAG	CTGGAAAGGT	TCIAAAGAAT	CCTTAGTTCC	TCAGAGGCAT								
K P G H V Q Q	S V N R K Q	S W K G	S K E S L V	P Q R H											
810	*	820	*	830	*	840	*	850	*	860	*	870	*	880	*
GGCCGCTAC	IAGGAGAAAG	IGTGGCCTAT	CATTCTGAGA	GTCCCACTC	ACAGACAGAT	GIAGGAAGAC	CTTTGCTGG								
G P P I G E S	V A Y H S E	S P N S	Q T D V G R	P L S G											
890	*	900	*	910	*	920	*	930	*	940	*	950	*	960	*
ATCTGGTAA	TCAGCAATIG	ITCAAGCTCA	CCCTAGCAAC	GGACAGAGAG	TGAACCCCCC	ACCACCACCT	CAAGTAAGGA								
S G I S A F	V Q A H	P S N G	Q R V N	P P P Q V R											
970	*	980	*	990	*	1000	*	1010	*	1020	*	1030	*	1040	*
GIGTACCTC	TCCACCACCT	CCAAGAGGCC	AGACTCCCCC	TCCAAGAGGT	ACAACITCCAC	CTCCCCCTTC	ATGGGAACCA								
S V I P P P P	P R G Q T P	P P R G	T T P P P P	S W E P											
1050	*	1060	*	1070	*	1080	*	1090	*	1100	*	1110	*	1120	*
AACTCTCATA	CAAGGGCTA	ITCTGGAAAC	ATGGAATACG	TATCTCCCG	AATCTCTCT	GTCCACCTG	GGGCATGGCA								
N S O I K R Y	S G N M E Y	V I S R	I S P V P P	G A W Q											

FIG. 9B

BAD ORIGINAL

30/43

1130	1140	1150	1160	1170	1180	1190	1200
AGAGGGCIAT	CCICCAACCAC	CTCTCAACAC	TTCCCCCATG	AATCCTCCTA	ATCAAGGACA	GAGAGGCATT	AGTTCIGTTC
E G Y P P P	P P P P L N T	S P M N P P	N Q G Q R G I	S S V			
1210	1220	1230	1240	1250	1260	1270	1280
*	*	*	*	*	*	*	*
CTGTGGCAG	ACAACCAATC	ATCATGCAGA	GTTCTAGCAA	ATTTAACITTT	CCATCAGGGA	GACCTGGAA	GCAGAAATGGT
P V G R Q P	I I M Q S S S K	F N F P S G	R P G M Q N G				
1290	1300	1310	1320	1330	1340	1350	1360
*	*	*	*	*	*	*	*
ACTGGACANA	CTGATTTCAT	GATACACCAA	AATGTTGICC	CTGCTGGCAC	TGTGAATCGG	CAGCCACCAC	CTCCATATCC
T G Q T D F	M I H Q N V V P	A G T V N R	Q P P P Y P				
1370	1380	1390	1400	1410	1420	1430	1440
*	*	*	*	*	*	*	*
TCTGACAGCA	GCTAATGGAC	AAAGCCCTTC	TGCITTIACAA	ACAGGGGGAT	CTGCTGCTCC	TTCGICATAT	ACAAATGGAA
L T A A N G	Q S P S A L Q T	G G S A A P	S S Y T N G				
1450	1460	1470	1480	1490	1500	1510	1520
*	*	*	*	*	*	*	*
GTATTCCTCA	GTCATGATG	GTGCCAAACA	GAAATAGTCA	TAACATGGAA	CTATATAACA	TTAGGTIACC	TGGACTGCAA
S I P Q S M	M V P N R N S H	N M E L Y N	I S V P G L Q				
1530	1540	1550	1560	1570	1580	1590	1600
*	*	*	*	*	*	*	*
ACAAATGGC	CICAGICAIC	TTCIGCTCCA	GCCCAGTCAT	CCCCGAGCAG	TGGGCATGAA	ATCCCTACAT	GGCAACCTAA
T N W P Q S	S S A P A Q S	S P S S G H E	I P T W Q P N				
1610	1620	1630	1640	1650	1660	1670	1680
*	*	*	*	*	*	*	*
CATACCAGIG	AGGICAAAT	CITITIAATA	CCCATTAGGA	AATAGAGCAA	GTACATCIGC	TAATTCICAG	CCTTCTGCTA
I P V R S N	S F N N P L G	N R A S H S	A N S Q P S A				

FIG. 9C

BAD ORIGINAL

31/43

1690	*	1700	*	1710	*	1720	*	1730	*	1740	*	1750	*	1760	*
CAACAGICAC	IGCAVHIACA	CCAGCTCCTA	TTCAACAGCC	IGTGAAAAGT	ATGCGTGTAT	TAAAACCAGA	GCTACAGACT								
T T V I A I I P A P I Q Q P V K S M R V L K P E L Q T															
1770	*	1780	*	1790	*	1800	*	1810	*	1820	*	1830	*	1840	*
GCTTACGAC	CIACACACCC	TTCTTGGATA	CCACAGCCAA	TTCAAACIGT	TCAACCCAGT	CCTTTTCCTG	AGGGAACCGC								
A L A P I I P S W I P Q P I Q T V Q P S P F P E G T A															
1850	*	1860	*	1870	*	1880	*	1890	*	1900	*	1910	*	1920	*
TTCAAAIGIG	ACIGIGATGC	CACCTGTTCG	TGAAGTCCCA	AACTATCAAG	GACCACCACC	ACCCTACCCA	AAACATCTGC								
S N V I V M P P V A E A P N Y Q G P P P Y P K H L															
1930	*	1940	*	1950	*	1960	*	1970	*	1980	*	1990	*	2000	*
TGCACCAAAA	CCCAICIGIT	CCTCCATACG	AGTCAATCAG	TAAGCCTAGC	AAAGAGGATC	AGCCAAGCIT	GCCCAAGGAA								
I H Q H P S V P P Y E S I S K P S K E D Q P S L P K E															
2010	*	2020	*	2030	*	2040	*	2050	*	2060	*	2070	*	2080	*
GATGAGAGIG	AAAGAGGIIA	IGMAAIGIT	GATAGIGGG	ATAAAGAAAA	GAAACAGATT	ACAACITCAC	CTATTACITG								
D F S I K S Y E N V D S G D K E K K Q I T T S P I T V															
2090	*	2100	*	2110	*	2120	*	2130	*	2140	*	2150	*	2160	*
TAGGMAAAC	AMGAMGAIG	AMGAGCGMAG	GGATCICIGT	ATTCAAGIT	ATTCTCTCA	AGCATTTAAA	TTCITTAIGG								
R K N K K D E E R R E S R I Q S Y S P Q A F K F F M															
2170	*	2180	*	2190	*	2200	*	2210	*	2220	*	2230	*	2240	*
AGCAATCIGT	AGMAAIGIA	CICMAATCTC	ATCAGCAGCG	TCTACATCGT	AAAAACAAT	TAGAGAATGA	AATGATCGCG								
I Q H Y I H V I Y S H Q Q R L H R K K Q L F N E M M R															

FIG. 9D

BAD ORIGINAL

32/43

2250	*	2260	*	2270	*	2280	*	2290	*	2300	*	2310	*	2320	*
GTTGGATTAI CTAAGAIGC CCAGGAICAA ATGAGAAAGA TGCITIGCCA AAAAGAAATCT AATTACATCC GTCTTAAAG															
V G L S Q D A Q D Q M R K M L C Q K E S N Y I R L K R															
2330	*	2340	*	2350	*	2360	*	2370	*	2380	*	2390	*	2400	*
GGCTAAMAG GACAAGTCTA IGTTTGIGAA GATAAGACA CTAGGAATAG GAGCAITTTGG TGAAGTCTGT CTAGCAAGAA															
A K M D K S M F V K I K T L G I G A F G E V C L A R															
2410	*	2420	*	2430	*	2440	*	2450	*	2460	*	2470	*	2480	*
AAGTAGATAC TAAGGCTTIG TAIGCAACAA AACTCTTCG AAAGAAAGAT GTTCTTCTTC GAAATCAAGT CGCICAIGTT															
K V D I K A L Y A T K T L R K K D V L L R N Q V A H V															
2490	*	2500	*	2510	*	2520	*	2530	*	2540	*	2550	*	2560	*
AAGGCTGAGA GAGATATCCT GGCIGAAGCT GACAATGAAT GGGTAGTTCG TCTATATAT TCATTCCAAG ATAAGGACAA															
K A E R D I L A E A D N E W V V R L Y Y S F Q D K D N															
2570	*	2580	*	2590	*	2600	*	2610	*	2620	*	2630	*	2640	*
TTTATACIII GAAAGGACI ACATTCCTGG GGGTGATAIG ATGAGCCTAT TAATTAGAAT GGGCATCTTT CCAGAAAGIC															
L Y I V M D Y I P G G D M M S L L I R M G I F P E S															
2650	*	2660	*	2670	*	2680	*	2690	*	2700	*	2710	*	2720	*
TGGCAGGAT CACATAGCA GAACCTACCT GTGCAGTTGA AAGTGTTCAT AAAATGGGT TTATTCATAG AGATATTAAA															
L A R I Y I A E L T C A V E S V H K M G F I H R D I K															
2730	*	2740	*	2750	*	2760	*	2770	*	2780	*	2790	*	2800	*
CCTGATAMIA ITTIGATIGA TCGIGAIGGT CATATTAAAT TGACIGACTT TGGCCICIGC ACTGGCTTCA GATGGACACA															
P D N I I I D R D G H I K L T D F G L C T G F R W T H															

FIG. 9E

BAD ORIGINAL



33/43

2810	*	2820	*	2830	*	2840	*	2850	*	2860	*	2870	*	2880	*
CGATCTCAG	IACIATCAGA	GIGGIGACCA	TCCACGGCAA	GATAGCATGG	ATTTCAGTAA	TGAATGGGGG	GATCCCTCAA								
D S K Y Y Q	S G D H	P R Q	D S M	D F S N	E W G	D P S									
2890	*	2900	*	2910	*	2920	*	2930	*	2940	*	2950	*	2960	*
GCTGTCGAG	TGGAGACAGA	CIGAAGCCAT	TAGAGCGGAG	AGCTGCACGC	CAGCACCAGC	GATGCTAGC	ACATTCTTGG								
S C R C G D R	L K P	L E R R	A A R	Q H Q	R C L A	H S L									
2970	*	2980	*	2990	*	3000	*	3010	*	3020	*	3030	*	3040	*
GTTGGGACTC	CCATTATAT	TGCACCTGAA	GTGTTGCTAC	GAACAGGATA	CACACAGTTG	TGTGATTGGT	GGAGTGTGG								
V G I P N Y I	A P E	V L L	R T G Y	T Q L	C D W	S V G									
3050	*	3060	*	3070	*	3080	*	3090	*	3100	*	3110	*	3120	*
TGTTATCTT	TTGAAAGT	TTGGGGACA	ACCTCCTTTC	TTGGCACA	CACCATTAGA	AACACAAATG	AAGTTATCA								
V I L I E M	L V G Q	P P F	L A Q	T P L E	T Q M	K V I									
3130	*	3140	*	3150	*	3160	*	3170	*	3180	*	3190	*	3200	*
ACTGGCMAAC	ATCCTCTTAC	ATCCACCAC	ANGCTAACT	CAGTCTGAA	GCTTCIGATC	TTATTATTAA	ACTTTGCCGA								
N W Q T S L H	I P P	Q A K L	S P E	A S D	L I I K	L C R									
3210	*	3220	*	3230	*	3240	*	3250	*	3260	*	3270	*	3280	*
GGACCCGAG	ATCGCTTAGG	CAAGAATGGT	GCTGATGAAA	TAAAGCTCA	TCCATTTTTT	AAAACAATTG	ACTTCTCCAG								
G P E D R L G	K N G	A D E	I K A H	P F F	K T I	D F S S									
3290	*	3300	*	3310	*	3320	*	3330	*	3340	*	3350	*	3360	*
TGACCTCAGA	CAGCAGCTC	CTTCATACAT	TCCTAAATC	ACACACCCAA	CAGATACATC	AAATTTTGAT	CCTGTTGATC								
D I R Q Q S	A S Y I	P K I	T H P	T D T	S N F	D P V D									

FIG. 9F

BAD ORIGINAL

34/43

3370	*	3380	*	3390	*	3400	*	3410	*	3420	*	3430	*	3440	*
CTGATAAAT		AICGAGIGAT		GATAACGAGG		AAGAAAATGT		AAATGACACT		CTCAATGGAT		GGTATAAAAA		TGGAAGCAT	
P D K I		W S D		D N E		E N V		N D T		L N G		W Y K		N G K	
3450	*	3460	*	3470	*	3480	*	3490	*	3500	*	3510	*	3520	*
CCTGACCAIG		CAICIAIGA		AITACCTTC		CGAAGGTTTT		TTGATGACAA		TGGCTACCCA		TATAATTAIC		CGAAGCCTAT	
P E H		A F Y E		F T F		R R F		F D D		N G Y		P Y N		Y P K	
3530	*	3540	*	3550	*	3560	*	3570	*	3580	*	3590	*	3600	*
TGAATAIGAA		IACATTAAT		CACAAGGCTC		AGAGCAGCAG		TCGGAIGAAG		ATGATCAAAA		CACAGGCICA		GAGATTAAAA	
E Y E		Y I N		S Q G		S E Q		S D E		D D Q		N T G		S E I	
3610	*	3620	*	3630	*	3640	*	3650	*	3660	*	3670	*	3680	*
ATCGCGATCI		AGIAIATGTT		TAACACACTA		GTAAATAAAT		GTAATGAGGA		TTTGTAAGAG		GGCCTGAAAT		GCGAGGIGTG	
N R D		L V Y V		*		*		*		*		*		*	
3690	*	3700	*	3710	*	3720	*	3730	*	3740	*	3750	*	3760	*
TTGAGGIIIC		GAGAGIAAAA		TIATGCAAT		ATGACAGAGC		TATATATGIG		TGCTCTGIGT		ACAAATATTTT		ATTTTCCTAA	
3770	*	3780	*	3790	*	3800	*	3810	*	3820	*	3830	*	3840	*
ATTAIGGGAA		AICCIHIIIA		AMIGTTAAT		TATTCAGCC		GTTTAAATCA		GTATTAGAA		AAAAATIGTT		ATAAGGAAAG	
3850	*	3860	*	3870	*	3880	*	3890	*	3900	*	3910	*	3920	*
TAAATTAIGA		ACIGAAATAT		ATAGTCAGTT		CTTGGIACCT		AAAGTACTTA		AAATAAGTAG		TGCTTIGTTT		AAAAGGAGAA	
3930	*	3940	*	3950	*	3960	*	3970	*	3980	*				
ACCIIGGATC		IATIGIATA		IAIGCTAAT		AATTTIIMAA		IACAAGAGIT		TTTGAAATTT		TTTT			

FIG. 9G

BAD ORIGINAL

35/43

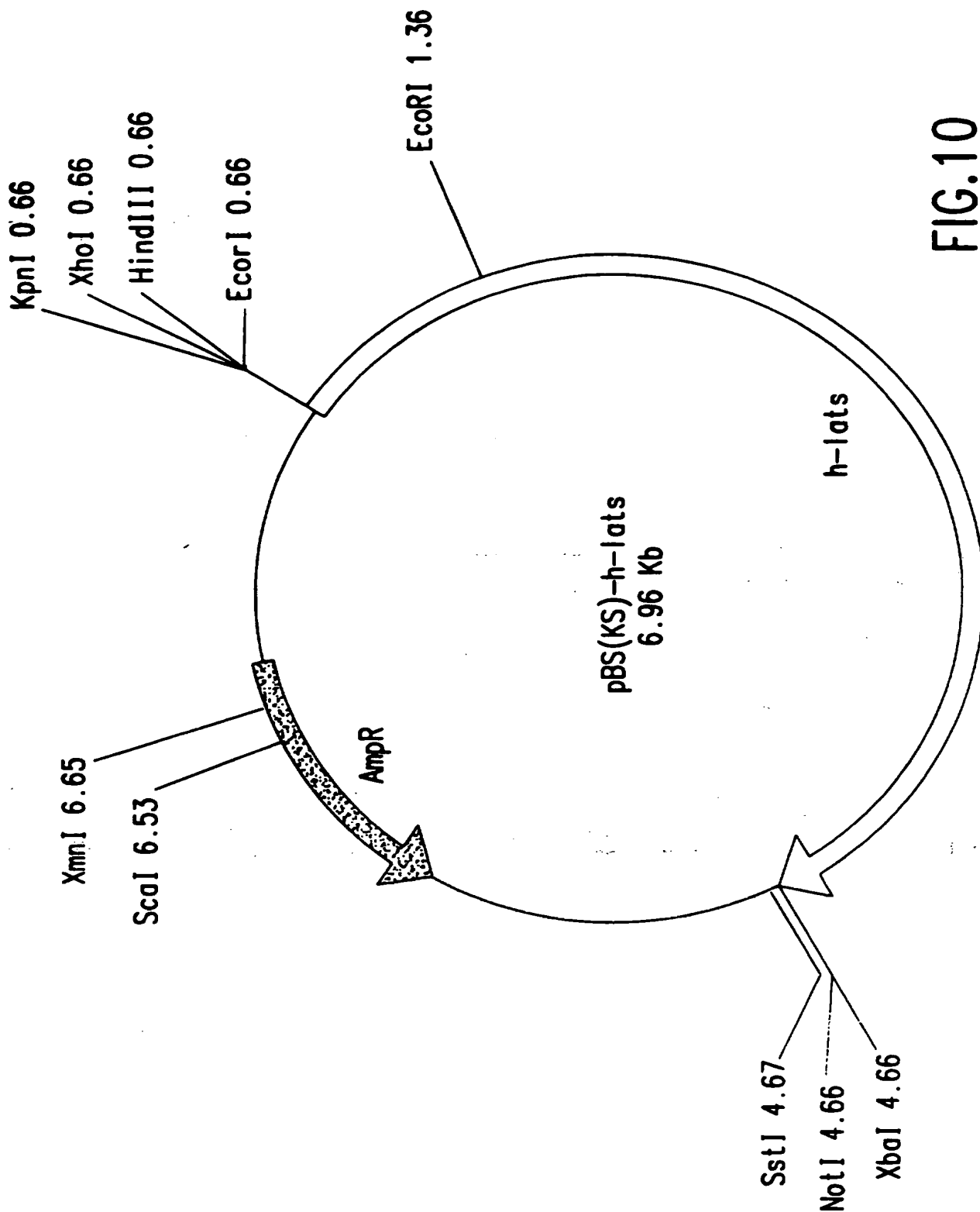


FIG.10

36/43

```

111 A1S MKRSEKPEGYRQMRPKTFPASNYTVSSRQMLQEIRESLRNL SKPSDAAKAEHNMSKMSTEDPRQVRNPPK 70
111 A1S IGTHIKAI QFIRNSLLPFANETNSSRSTSEVNPQMLQDLQAAGFEDMWIQALQRTNNRSIEAAIEFISK 140
111 A1S HSYQDPRREQMAAAAAARPINASMKPGNVQQSVNRKQSWKGSKESLVPQRHGPPLGESVAYHSESPNSQTD 210
111 A1S .h.i.....s.n.v.r.....a. 45
111 A1S VGRPLSGSGISAFVQAIIPSNQGRVNP PPPPQVRSVTPPPPRGQTPPPRGTTTPPPPSWEPNSQTKRYSGN 280
111 A1S .a..a.....s..... 117
111 A1S HIIYVISRISPVPPGAWQEGYPPPLNTSPMNPNGQGRGISSVPVGRQPIIMQSSSRFNFPGRPGMQNG 350
111 A1S .t.....s.a.a.....t.....tp.....v... 187
111 A1S IGQIDFMIIHQNVVPAGTVNRQPPPPYPLTAANGQSPSALQTGGAAPSSSYNGSIPQSMVMPNRNSHME 420
111 A1S .q.s.iv...t.s.t.....p.....a...p.fa.nv..... 256
111 A1S IYNI SVYGIQINMPQSSSAPAQSSPSSGIEIPTWQNPVRSNFSNNPLGNRASHSANSQPSATTIVTAIT 490
111 A1S .n.....a.....g.....s..... 326
111 A1S PAPIQQPVKSMRVLKPELQIALAP'HIIPSWIPQPIQTVQSPFPEGTASNVTVMPPVAEAPNYQGPPPPYP 560
111 A1S .....m...v.....t.s.....s.p.i.....s..... 396
111 A1S KIILLHQNPVPPYESISKPSKEDQPSLPKEDESEKSYENVDSGDKEKKQITTSPIITVRKNKKDEERESR 630
111 A1S .....v...c.de.....d....adsg..... 466
111 A1S IQSYSPQAI'K'FTFMEQIVENVLKS HQRLHRKKOLENEMMRVGLSQDAQDQMRKMLCQKESNYIRLKRK 700
111 A1S ..... 536

```

FIG. 11A

BAD ORIGINAL

37/43

hLATS	DKSFI VKIKIIGIGATGIVCLARKVDTKALYATKTLRKDVLLRNQVAHVKAERDILAEADNEWVRLYY	770
mLATS	.....	606
hLATS	SIQJIKINLYIVMDYIPGGDMMSLLIRMGIFPESLARFYIAELTCAVESVHKMGFIHRDIKPDNIIJDRDG	840
mLATS	.....n.....	676
hLATS	IIIKLIDFGLCTGFRWTHDSKYYQSGDHPQDSMDFSNWGDPSRCRGDRLKPLERRAARQHRCIAHSL	910
mLATS	.....n.....	746
hLATS	VGITPNYIAPEVILLRIGYTQLCDWMSVGVILFEMLVGQPPFLAQTPLETQMKVINWQTSIHPPQAKLSPE	980
mLATS	.....c.....i.....	816
hLATS	ASPIIITICRGPEDRLGKNGADEIKAIIPFFKTIIDFSSDLRQQSASYIPKITHPTDTSNFDPPVDPDKLWSD	1050
mLATS	.....	886
hLATS	DNIIIVNIIHNGWYKNGKIPEHAFYEFTFRITDDNGYPYNYPKPIEYEYINSQSEQSDDEDDQNTGS	1120
mLATS	qss...iss...s.....h.....h.s.	
hLATS	LTIKNIRDI VYV	1130
mLATS	dgn.....	966

FIG. 11B

BAD ORIGINAL

38/43

```

hlA1S MKRSEKPEGYRMRPKTFPASNYTVSSRQMLQEIRESLRNL SKPSDAAKAEHNMSKMSTEDPRQVRNPPK 70
mlA1S? m.at... 45

hlA1S IGIHIIKALQEI RNSLLPFANETNSSRSTSEVNPQMLQDLQAAGFDEDMV IQALQKTMNRSIEAAIEFISK 140
mlA1S? .pyq...r...y...sgt...aaa...r...e.vn.ac.qe.agr..tq.gs.....y... 114

150 160 170 180 190 200 210
hlA1S MSYQDIPRREQMAAAARPINASMKPGNVQSVNRKQSWKGSKESLVPQRHGPPLGESVAYHSESPNSQTD 210
mlA1S? .q.l...n..i-vrvikqtspg-..lastp.t.rp.fe.tg.a.-.sy.--q...-gan...-g.aalee 175

hlA1S VGRIP'LSGSGISAFVQAHPSTNGQVRNPPPPQVRSVTPPPPRGQTPPPRGTTPPPPSWEPSQTKRYSGN 280
mlA1S? mmp.qy-----ldf1fpgag

hlA1S MLYVVISRISPVPPGAWQEGYPPPPPLNISPMPNPQNGQRISSVPVGRQPIIMQSSSKFNIPSGRPGMQNG 350
mlA1S? aqfhaqaqahq...-k...-stave.sahfpgthy.rghl1seqsgyv.r..s.q-nktp.dayss 251

hlA1S IGIIDIMIIQNVPAGTVNRQPPPPYPLIAANGQSPSALQTIGGSAAPSSYTNGSIPQSMMPNPNRNSHME 420
mlA1S? mak.aqqgppaslt.spahaglytashhk-p...tppgahp.hv1.trg.-tf.ge.sa.avla.s...l.ad 319

hlA1S IYNISVPGLQINWPQSSSAPAQSSPSSGHEIPTWQPNIPVRSNSFNPNPLGNRASHSANSQPSATTVTAIT 490
mlA1S? .elq-stv--p.saapl.frd.lqkq...-asr-.hvaf.agp-srtnsfnnpqpep.l.apn....v. 383

hlA1S PAP'IQPVKSMRVLKPELQITALAPTHPSWIPQPIQTVQSPFPEGIASNVTMPPPVAEAPNVQGPPPPYP 560
mlA1S? a.h.lh...v...r..pQ...vg.s..a.vaa.tapate.letkegsagphpldvdyggserrc..... 453

hlA1S RHHIHQNP'SVP---PYESISKPSKEDQPSL'PKEDESEKSYENVDSGDKEKKQITTSPTIVRKNKKDEERRESR 630
mlA1S? ....lpSk.eqySvdld.lCtsvqqslrggtdl.g.d.hakg.kagrd....q...vp....sr...k.... 528

```

FIG. 12A

BAD ORIGINAL

39/43

```

hlAIS IQSYSPQAFKFFMEQHVENVLKSHQQLHRKKQLENEMRVGLSDAQDQMRKMLCKESNYIRLKRKM 700
mlAIS? .k...y.....i.ty.kvs.rl...q.aka.ceae.e...i.y.....n..... 598

hlAIS DKSHFVKIKTLGIGAFGEVCLARKVDTKALYATKTLRKKDVLRLNQVAHVXAERDILAEADNEWVRLYY 770
mlAIS? .....c.l.h...m.....n.....k.... 668

hlAIS SIQIKDNI.YFVMDYIPGGDMMSLLIRMGIFPESLARFYIAELTCAVESVHKMGFIHRDIKPDNILIDRDG 840
mlAIS? .....s.....ev.h.....l.i.....l... 738

hlAIS IHHKIIDI GLCTGFRWTHDSKYYQ-SGDHPRQDSMDFSNEWGDPSSCRCGDRLKPLERRAARQHQRCLAHSL 910
mlAIS? .....n.....-k.n.m.....epgd.l.d.v.n.....t..q..qk..... 808

hlAIS VGIPINYIAPEVLLRTGYTQLCDMWSVGVILFEMLVGQPPFLAQTPLETQMKVINWQTSLSLHIPPQAKLSPE 980
mlAIS? .....k.....p.t...l.....est....t.vr..a. 878

hlAIS ASHIIIFI CRGPIIDRLGKNGADEIKAHPPFKTIDFSSDLRQQSASYIPKITHPTDTSNFDVPDPDCLWSD 1050
mlAIS? .....rd...dl.....n.....r.i.k.a.p.v.t.s.m.....eesp.he 948

hlAIS DHHIILHVN-DIINGWYKNGKIIPETIAFYEFTHRIIDNGYPYNYPKPIEYEYINSQGSEQSDEDDQNTGS 1120
mlAIS? asq.sakaw...as--pss.....frc...s.paesadpgdadleg ----- 1009

hlAIS IIEHRLVYV 1130
mlAIS? aaegcqp... 1119

```

FIG. 12B

BAD ORIGINAL

40/43

		<b>LSD2a</b>		
h-LATS	MKRSEKPEGYROMRPKTFPASNYTVSSROMLQE IRESLRNLSKPSDAAKAEHNMSTMSTEDPROVRNPPK-	70		
LATS	_____ Mh. agekrgrnd.yta.alesikqdltr	30		
h-LATS	FGTHHKALQEIRNSLLPFANEINSSRSTSEVNQMLQDLQAAGFDEDMV IQALOKTNNRSIEAAIEFISK	140		
LATS	.evqnnhr:nnq=.ytp.ryta..grndaltptyhhakqpmepppsaspapdvv-ippppa.vgqpgag.-	97		
h-LATS	MSYQDPRREQMAAAAAARPINAMKPGNVQOSVNRKQSWKGSKE SLVPQRHGPPLGESVAYHSE-SPNSQTD	210		
LATS	i.vsgvgvgvgvgv.ng.-v-p-.mtalmpnkli..p.ierdta.shyl.cs.a.dsgagssrsd..h.h-h	165		
		<b>SH3-BINDING</b>		
h-LATS	VGRPLSGSGISAFVQAHPNGORVNP PPPPVRSVT PPPPPRGQT PPPRGTT PPPPSWE PNSQTKRYSGN	280		
LATS	thq.----s.r.t.gnpgg..g-fs.s.sgfsevap.a....np.assaa.p...vpplsqayv..r.po	229		
		<b>LSD1a</b>		
h-LATS	MEYVISRISPVPAGAWQEGYPPPLNTSPMPPNQGQGISSVPVGRPIIMQSSSKFNFPSCRPGMONG	350		
LATS	Innrppa.a.ptqrgnspvitqng.k-n.qqqll.qikslny.g.qsgavvepppyliqg.ag.aapp	298		
h-LATS	TGQTD FMIHQNVVPAGTVNRQPPPPYPLTAANGQSPSALQTGCSAAPSSYTNGSIPQSMV PNRNSHME	420		
LATS	pppsytasmqsrqsp.qsq.s--d.rkspss.iy--tsa..ps.itvsippa.lakpq.rvyqarsq	364		
h-LATS	LYNISVPG LQTNMPQSSSA--PAQSSPSSGHEIPTWQPNIPVRSNSFNNPLGNRASHANSQPSATTVTAIT	490		
LATS	qpi.mqsvks.qvqkpvltav.pq...asasnsphvlsappsyppksoavvqqqqqaaaaahqqhqhq	436		
		<b>LSD1a</b>	<b>LSD1p</b>	<b>LSD2a</b> <b>LSD2p</b>
h-LATS	PAPIQQPVKSMRVLKPELQTALAPTHPSWIPOPIQTVQPSPFEGTASNVTMPPVAEAPNYQGP PPPPYp	560		
LATS	qskppt.ttppl.glnskpnc.e.psyaksm.akaatvv erdqrererdqaklangnpgq qml.....q	545		
		qqqqqqqqqqvqqqqvqqqqqqqqqqlqalrvlqqaaqr		snnnnnseikppscnnnni

FIG.13A



41/43

LFD

h-LATS KHLHQNPSVPYESSKPSKEDQPSLPKEDESEKS-YENVDSGDKEKKQITTSPI TVRKN-K-KDEERRESR 630  
 LATS qisnsnlatt..ipvkynnnssntganssgg.ng.tgttas.stsc..ikho...pe..kis.e.e...k.f. 638

h-LATS IQSYSPQAFKFFMEQHVENVLKSHQORLHRKKQLENNMRVGLSQDAQDQMRKMLCQKESNYIRLKRAKM 700  
 LATS .rq.....i....i..yr..ty..n...k..hk...pdgt.ie.....n..... 708

LFD — KINASE DOMAIN

h-LATS DKSMFVKIKTLGIGAFGEVCLARKVDT-KALYATKTLRKKDVLRLNQVAHVKAERDILAEADNEWVRLYY 770  
 LATS .....pi.v.....l.vs.i..snh...m.....a...k.....n...k... 779

h-LATS SFQDKONLYFVMDYIPGGDMMSLLIRMGIFPESLARFYIAELTCAVESVHKMGFIHRDIKPDNILIDRDG 840  
 LATS .....l.....kl...e.e.....v....d..... 849

h-LATS HIKLTDFGLCTGFRWTHDSKYYQ-SGDHPRQDSMDFSNEWGDPSSCRCDRLKPLERRAARQHQRCLAHS 910  
 LATS .....n.....en.n.s.....e-p—eey.e-n-.pkptv....rm.d...v.... 915

h-LATS VGTPNYIAPEVLLRTGYTQLCKWWSVGVILFEMLVGQPPFLAQTPLETQMKVINWQTSLHIPPQAKLSPE 980  
 LATS .....e.s.....y.....y.....ns.....q.....ekt.....e..r. 985

KINASE DOMAIN

h-LATS ASDLI IKLCRGPEDRLGKNGADE IKAHPFFKTIDFSSDLRQQSASYIPKITHPTDTSNFDPPDKLWSD 1050  
 LATS .t...rr...asodk....-sv..v.s.d...g...-a.m.k.k.p...e.k.....e..r.n 1053

h-LATS DNEEENVNDTLNGWYKNGKHPEHAFYEFTRRFDDNGYPYNPKPIEYIYINSQSEQQSDEDDQNTGS 1120  
 LATS .stmssgd.-vd---q.dr-tf.g.f.....kqp.dmt-----ap-- 1096

LCD2

h-LATS EIKNRDLVVY 1130  
 LATS ----- 1099

LCD3

FIG.13B

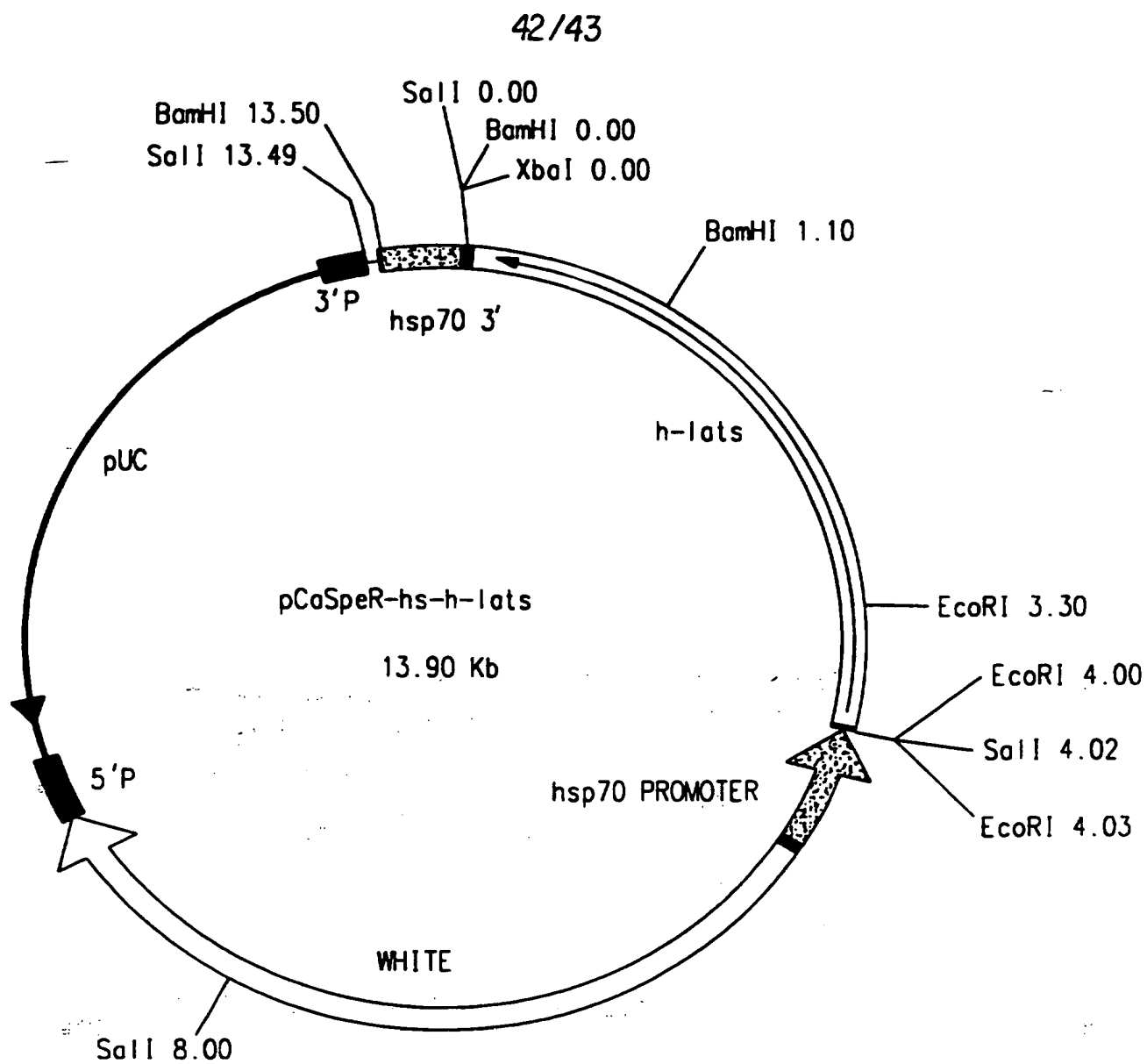


FIG.14

43/43

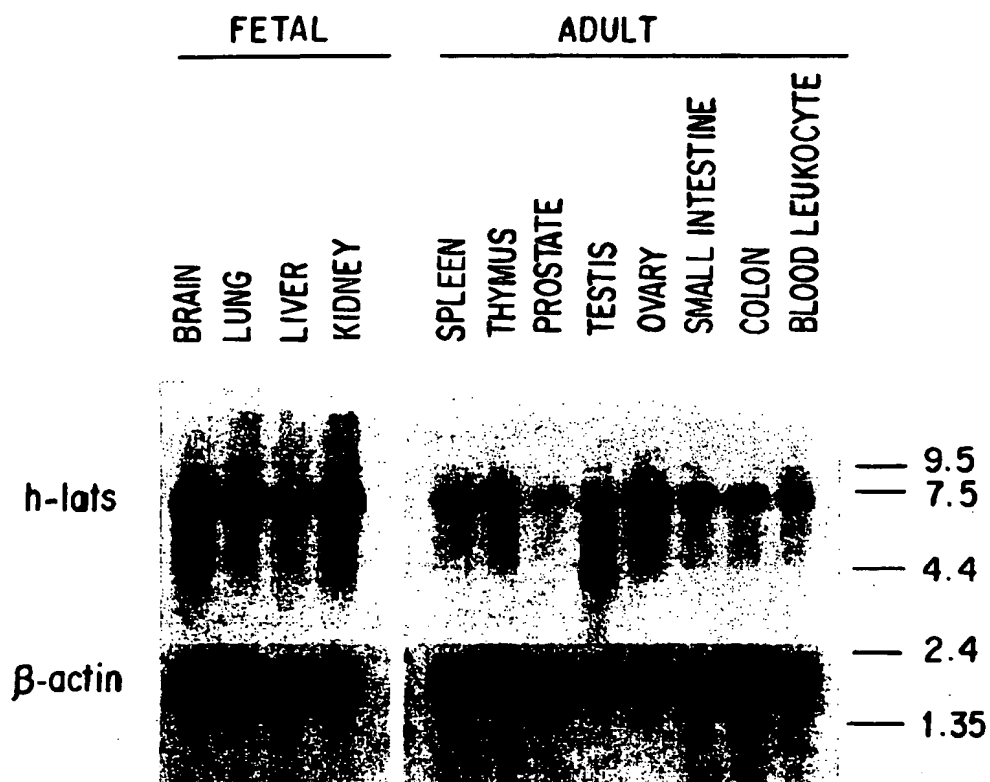


FIG.15

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US96/04101

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) :C07K 11/00; C07H 21/04; C12P 21/02; C12N 5/10; A61K 38/43

US CL :530/350; 536/23.2; 435/69.1, 240.1; 514/2

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 536/23.2, 23.4; 435/69.1, 69.7, 240.1; 514/2; 935/9

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, BIOSIS, IntelliGenetics

search terms: lats gene, drosophila tumor suppressor gene, sequence, nucleic acid, nucleotide, clone, expression

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ---- Y	GENES AND DEVELOPMENT, Volume 9, Number 5, issued 01 March 1995, Justice et al, "The <i>Drosophila</i> Tumor Suppressor Gene <i>warts</i> Encodes a Homolog of Human Myotonic Dystrophy Kinase and is Required for the Control of Cell Shape and Proliferation", pages 534-546, see entire document.	1, 3, 6-8, 10-19, 23-25, 28-32, 34-36, 38-39 ----- 2, 4-5, 9, 26-27, 33, 37, 40-52, 78
X	EMBO JOURNAL, Volume 11, Number 6, issued June 1992, Yarden et al, " <i>cot-1</i> , a Gene Required for Hyphal Elongation in <i>Neurospora crassa</i> , Encodes a Protein Kinase", pages 2159-2166, see entire document.	7-8, 10-11, 14-19, 28-30, 32, 35, 39



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	
*A* document defining the general state of the art which is not considered to be of particular relevance	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*E* earlier document published on or after the international filing date	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*O* document referring to an oral disclosure, use, exhibition or other means	*Z* document member of the same patent family
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

09 JULY 1996

Date of mailing of the international search report

25 JUL 1996

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

ROBERT C. HAYES

Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

 International application No.  
 PCT/US96/04101

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	THE JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 266, Number 19, issued 05 July 1991, Shortridge et al, "A <i>Drosophila</i> Phospholipase C Gene that is Expressed in the Central Nervous System", pages 12474-12480, see entire document.	7-8, 10-11, 14-15, 17-19, 28-30, 32, 35, 39
X	GENE, Volume 104, Number 1, issued 1991, Toyn et al, "The Cell-Cycle-Regulated Budding Yeast Gene <i>DBF2</i> , Encoding a Putative Protein Kinase, has a Homologue that is Not Under Cell-Cycle Control", pages 63-70, see entire document.	7-8, 10-11, 14-15, 17-19, 28-30, 32, 35, 39
X, P ---- Y, P	DEVELOPMENT, Volume 121, Number 4, issued April 1995, Xu et al, "Identifying Tumor Suppressors in Genetic Mosaics: the <i>Drosophila lats</i> Gene Encodes a Putative Protein Kinase", pages 1053-1063, see entire document.	1, 3, 6-8, 10-19, 23-25, 28-32, 34-36, 38-39 ----- 2, 4-5, 9, 26-27, 33, 37, 40-52, 78

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US96/04101

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-19, 23-52, and 78

Remark on Protest

☐

The additional search fees were accompanied by the applicant's protest.

☐

No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US96/04101

unity of invention is lacking.

Groups V and XIV contain claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention for the above reasons, which explain why the compositions used lack unity and are not so linked as to form a single inventive concept under PCT Rule 13.1. If the fee for searching Groups V or XIV is paid, the first named embodiment, the anti-lats antibody, will be searched. In order for more than one species to be examined, the appropriate additional examination fees must be paid. The species for claims 66-67, 69, 100-103 are as follows:

- A) anti-lats antibody.
- B) lats derivative or analog.
- C) lats antisense nucleic acid.
- D) a nucleic acid comprising a portion of the lats gene.

In Group V, the following claims are generic: claims 66-67, 69.

In Group XIV, the following claims are generic: claims 100-103.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US96/04101

### BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-19, 23-52 and 78, drawn to a purified lats protein, derivative, analog, or fragment, a chimeric protein, an isolated nucleic acid, a recombinant cell, a method of producing the lats protein and a pharmaceutical composition and a kit that comprises a lats protein.

Group II, claims 20-22, 56-57 and 77, drawn to an antibody, a molecule comprising antibody fragments and a pharmaceutical composition and a kit comprising these antibodies/fragments.

Group III, claims 53-55, 70-71 and 77, drawn to pharmaceutical compositions comprising a therapeutic nucleic acid, an oligonucleotide, a recombinant cell and a kit comprising the nucleic acid probes/primers.

Group IV, claims 58-65, drawn to a method of treating a disease state by administering a molecule that promotes lats function.

Group V, claims 66-69, drawn to a method of treating a disease state by administering a molecule that inhibits lats function.

Group VI, claim 72, drawn to a method of inhibiting expression of a nucleic acid with an oligonucleotide.

Group VII, claims 73-76, drawn to a method of diagnosis of a disease by screening aberrant levels of lats RNA or protein using nucleic acids or proteins or antibodies.

Group VIII, claims 79-80, drawn to a method to increase cell growth in plants.

Group IX, claims 79 and 81, drawn to a method to increase cell growth in animals.

Group X, claim 82, drawn to a method of screening for lats ligands.

Group XI, claims 83-85, drawn to transgenic plants.

Group XII, claims 83, 85, 92-95 and 99, drawn to transgenic animals and method of making.

Group XIII, claims 86-91 and 96-98, drawn to a method of identifying a tumor suppressor gene.

Group XIV, claims 100-103, drawn to a method of inhibiting cellular senescence in a subject.

Claim 77 has been placed in both Groups II and III. The antibody embodiment will be searched with Group II. The nucleic acid embodiment will be searched with Group III.

Claim 79 has been placed in both Groups VIII and IX. The plant embodiment will be searched with Group VIII. The animal embodiment will be searched with Group IX.

Claims 83 and 85 have been placed in both Groups XI and XII. The plant embodiment will be searched with Group XI. The animal embodiment will be searched with Group XII.

The inventions listed as Groups I-XIV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Group I is directed to purified lats protein, analogs, fragments, chimeric constructs, to the DNA that encode them and to a pharmaceutical composition and kit, which is the first appearing product, method of making and method of using. The special technical feature is the disclosed protein and DNA sequences. Group(s) II-III, XI-XII are drawn to structurally different products which do not share the same or a corresponding technical feature. Group(s) IV-X and XIII-XIV are drawn to methods having different goals, method steps and starting materials, which do not share the same or a corresponding special technical feature. Note that PCT Rule 13 does not provide for multiple products or methods within a single application. Since the special technical feature of the Group I invention is not present in the Group II-XIV claims, and the special technical features of the Group II-XIV inventions are not present in the Group I claims,